



**TECHNICAL UNIVERSITY OF KOŠICE**  
Faculty of Electrical Engineering and Informatics



# Electrical Engineering and Informatics V

**Proceedings of  
the Faculty of Electrical Engineering and Informatics  
of the Technical University of Košice**

**ISBN 978-80-553-1704-5**

**Electrical Engineering and Informatics V**  
**Proceedings of the Faculty of Electrical Engineering and Informatics**  
**of the Technical University of Košice**

Published by: **Faculty of Electrical Engineering and Informatics,  
Technical University of Košice, Letná 9,  
042 00 Košice, Slovak Republic**

Date of publication: **20.09.2014**      Language: **English, Slovak**

Printing: **50 pieces CD**      Pages: **994**

Editorial board chairman: **Prof. Ing. Dušan Kocur, CSc.**

Proceeding Reviewers: **Prof. Ing. Dušan Kocur, CSc.**  
**Prof. Ing. Alena Pietriková, CSc.**  
**Prof. Ing. Peter Sinčák, CSc.**  
**Prof. Ing. Dušan Levický, CSc.**  
**Prof. Ing. Michal Kolcun, PhD.**  
**Prof. Ing. Dobroslav Kováč, CSc.**  
**Prof. Ing. Daniela Perduková, PhD.**  
**Prof. Ing. Ján Sarnovský, CSc.**  
**Prof. Ing. Jozef Juhár, CSc.**  
**Assoc. Prof. RNDr. Dušan Olčák, CSc.**  
**Assoc. Prof. RNDr. Marián Klešč, PhD.**  
**Assoc. Prof. Ing. Jaroslav Porubän, PhD.**  
**Ing. František Babič, PhD.**

Technical Editor: **Ing. Ivan Halupka**



Optimal State Control of the Mechatronical Laboratory Model B&P_KYB .....	9
Principles of polymorphic types.....	17
Knižnica funkcií pre testovanie AD prevodníkov v LabVIEW .....	23
Fuzzy riadenie SMPM.....	29
Modul pre vybrané funkcie a zariadenia fiktívnej budovy ovládané prostredníctvom Internetu.....	34
Photorealistic images execution in parallel environment .....	40
Lygon - a programming language for linear logic .....	45
Analýza trhu s elektrinou.....	49
Meranie vlastností nelineárnych dielektrík .....	55
Application for 3D morphing .....	60
Prehľad využitia metód dolovania v dátach vo výrobnom priemysle .....	65
Možnosti využívania nízopotenciálneho tepla .....	70
Meranie vlastností solárnych systémov polohovaných smerom k Slnku .....	75
Výbojové javy v atmosfére.....	80
Digitálna Vodotlač v Transformačnej Oblasti vo Videu .....	86
Digitálna vodotlač vo videu .....	91
Pattern recognition in motion with sensors of mobile devices.....	96
Global electrical power planning, investments and projects .....	100
Design of Secure Server using the Micro Kernel and verification.....	105
Bezdrôtový prenosový systém s viacerými vstupmi a viacerými výstupmi .....	110
Improving Reliability of Arena Platform for Automated Assessments .....	115
Data collection and evaluation for the purpose of assistive applications for road transport .....	119
Basics of formal specification of component programming.....	124
Integration of new ICT solutions for Self-Government.....	130
Ovládanie robotického ramena pomocou dátovej rukavice .....	136
Odporúčací systém vytvorený v prostredí Apache Mahout .....	141
Využitie DCT a SVD pri návrhoch algoritmov digitálnej vodotlače .....	147
Ukrývanie dát vo videosekvenciách na báze CDMA .....	153
Steganografia s Využitím Retransmisie Segmentov v Protokole TCP .....	157
Moderný Prístup Obrazovej Steganalýzy .....	163
Position Controller for Permanent Magnet DC Machine.....	169
Utilisation of network simulators for teaching computer network oriented subjects.....	175
Security extension for live streaming in IP networks.....	180
Využitie platformy Arduino vo fotovoltike.....	185

WebRTC implementation for simple videoconference solution .....	191
EIGRP protocol analysis.....	195
Extension of smartphone functionality using external sensors.....	201
Teplotná analýza elektrických kontaktov vysokoprúdových zariadení.....	206
Application for monitoring of the CNL Videosever platform .....	210
Real time object tracking and streaming Android application .....	214
Conditional video streaming from mobile devices in multi-cameras environment .....	218
Eliminácia vplyvu prechodných dejov na elektrickú sieť .....	222
Vplyv parametrov kontaktovania na tvar slučky kontaktovaného spoja .....	228
Analýza dynamických dejov v elektrizačnej sústave využitím SimPowerSystems .....	233
Lakovací robot.....	239
Elektronické riadenie stability vozidla .....	243
Riadenie veternej elektrárne .....	248
Simulácia ASR systému .....	252
Využitie vizualizácie pri výučbe algoritmov a údajových štruktúr .....	257
Verifying Secrecy and Authenticity of Security Protocols.....	262
OrCAD PSpice – model výkonového MOSFET tranzistora .....	267
Comparison of Feature Extraction Algorithms for the Acoustic Event Detection System .....	273
Inteligentná elektrická inštalácia CyBrotech .....	279
Inteligentný elektroinštalačný systém BTicino .....	284
Elektrická požiarňa signalizácia obchodného priestoru .....	288
Design of software tool for energy certification of buildings .....	292
Prenos informácií v elektrických staniach.....	297
Diskriminačná analýza reči založená na adaptívnom modelovaní foném .....	300
3D Editor and Simulator of Petri Nets.....	306
Application of forensic tools to gain description of software behavior .....	311
Automatizácia generovania požiadavky na zmenu v procese manažmentu zmien .....	316
Interaktívna vizualizácia zhukovania textov použitím algoritmu GHSOM v cloudovom prostredí .....	320
Optimalizácia distribúcie algoritmu pre tvorbu klasifikačných stromov pre úlohy dolovania v textoch ..	324
Miniaturizovaný systém pre podporu rehabilitácie dolnej končatiny .....	329
Rozpoznávanie hovoriaceho z emocionálnej reči na princípe i-vektorov.....	333
Prototyp laboratórneho zariadenia pre sušenie a vytvrdzovanie hrubých vrstiev .....	339
Vývoj výučbového zariadenia pre subtraktívnu prípravu DPS .....	343
Prieskumná Analýza Publikačnej Činnosti .....	348

Stereoskopické meranie vzdialenosti v automatickom inventarizačnom systéme dopravných značiek ..	351
Optimalizácia Internetového Predaja Pomocou Vhodnej Marketingovej Stratégie .....	356
Inventory System for Trafic Signs .....	361
Mobilná Aplikácia Využívajúca Prvky Umelej Inteligencie .....	367
Efektívnejšie Využívanie Výrobného Odpadu Prostredníctvom Jednoduchej Aplikácie .....	371
Klasifikácia Emailov Pomocou Vhodných Metód Dolovania v Dátach.....	375
On-line Aplikácia pre Komplexný Manažment Taxislužby .....	380
Riadenie pre autonómne hydroponické systémy .....	384
Analýza vložených páuz a dysfluentných javov v spontánnom rečovom prejave .....	389
Identifikácia autorít v konverzačnom obsahu .....	395
Limitations of Dispersion and Pulse Broadening in Optical Fiber .....	399
Chránenie malých energetických zdrojov .....	405
Overovanie činnosti ochrán transformátorov .....	410
Chránenie trakčných napájacích staníc .....	416
Overovanie činnosti ochrán proti zemným spojeniam .....	421
Odporúčací Systém nad Databázou Filmov .....	426
Metódy strojového učenia v analýze názorov .....	431
Advanced organization of Bibliography and Citations.....	435
Odhad Času 3D Modelovania s Využitím Petriho Sietí .....	441
HMM syntéza reči slovenského jazyka s využitím adaptácie modelov .....	445
Design and Measurements of Microstrip Narrow-band Notch Filter Based on LTCC Substrate Murata LFC.....	451
Reasons to Use Formal Methods in Many Areas .....	455
Kategorizácia textových dokumentov v slovenskom jazyku metódami kontrolovaného učenia .....	461
Vzdialené pracovisko pre FPGA obvody Altera.....	465
Viackanálový prenos dát vo frekvenčnom pásme SRD .....	469
Detekcia, sledovanie a identifikácia pohyblivých cieľov vo videosekvenciách s využitím časticových filtrov.....	473
Universal model of operating systems migration.....	477
Virtualizovaná laboratórna architektúra s využitím platformy od spoločnosti Microsoft .....	483
Použitie evolučného algoritmu pre optimalizáciu energie .....	488
Meranie účinnosti tienenia a odrazu elektromagnetického poľa pre stavebný objekt polystyrén .....	496
Analýza procesných dát z pohľadu organizačnej perspektívy .....	502
Podpora vybraných fáz dolovania procesov v Matlabe .....	506

Analýza, modelovanie a simulácia skladových procesov.....	510
Prepínace v Hybridných RF/FSO Komunikacných Systémoch.....	515
Hybridné RF/FSO Komunikacné Systémy.....	519
Vyhľadávanie podozrivých tém v sociálnom webe prostriedkami strojového učenia .....	525
Detekcia typu prehovoru používateľa.....	530
Solving Economical Problems by Vector Optimization Methods .....	535
Simulácia teploty výkonových prvkov v elektronických štruktúrach s využitím softvéru HyperLynx® Thermal.....	542
Prototyp palubného počítača na báze OBD2 s bezdrôtovým rozhraním .....	548
Design of a Wireless Sensor Network for the Surface Water Quality Monitoring .....	552
Vplyv miesta pripojenia alternatívneho zdroja energie na veľkosť skratového prúdu v distribučnej sústave .....	558
Secured web oriented videoconference system .....	563
8-bit ATMELE microcontroller based Street view bicycle.....	568
Development of automated methods for network traffic filtering depending on a client station behavior .....	572
Environment for creation and reviewing of learning materials in the form of LTI application for Canvas system.....	577
Man-Machine rozhrania pre mobilné zariadenia v IaRS.....	581
Kontrolný video dohľadový systém pre monitorovanie cestnej premávky v tuneli Branisko .....	585
Automated System for Operations Management and Crisis Events for Smart Home Application .....	591
Web application of smart home .....	597
Control module for smart house based on Arduino and its modules .....	602
Aplikácia slovníkového prístupu pri analýze sentiment.....	606
Testovanie moderných steganografických metód.....	613
SDI-12 adaptér pre monitorovanie dát zo senzorov kvality vody.....	618
Predspracovanie dát z UWB radaru pomocou Mini PC MK802 III.....	623
FIR Filter Implementation for eDiViDe Remote Laboratory .....	628
A Low Dropout P-Channel MOSFET based GPS Backup Circuit .....	633
Klasifikácia diskusných príspevkov pomocou metód strojového učenia .....	637
Identifikácia autorstva dolovaním obsahu v sociálnom webe.....	641
PWM Modulator for Multilevel Inverter.....	645
Automatická segmentácia reči v programe Praat .....	650
Impulzný ohrev, ako možnosť úpravy magnetických vlastností feromagnetických amorfných zliatin .....	654
Methods of student knowledge visualization.....	659

Klasifikácia názorov dolovaných v konverzacnom obsahu .....	663
Kodér a dekodér stavového binárneho aritmetického kódovania .....	668
Rozpoznávanie sarkastických viet v konverzačnom obsahu .....	674
Porovnanie metód transformácie kardinalitných ohraničení na SAT problémy .....	678
Evolutionary agent model creation .....	683
Analýza mechanických vlastností LTCC substrátov .....	688
Design of Low Pass Filter Based on Substrate Murata LFC for Ultra Wide-Band Sensor System .....	693
Simulation of a Two-Phase Induction Motor .....	698
Učenie sa chôdze mobilného robota neurónovou sieťou s využitím metódy ZMP .....	708
Fuzzy logika pri detekcii poisťovacích podvodov .....	714
Fuzzy kognitívne mapy v robotických úlohách kopu .....	719
Simulation of the Vehicle Navigation Using Fuzzy Cognitive Maps .....	725
Adaptácia fuzzy kognitívnych máp pomocou algoritmu PSO v robotickom futbale.....	729
Využitie fuzzy sviečkových grafov pri vyhodnocovaní ekonomických dát .....	735
Energetická Optimalizácia Regulátora Dynamického Systému.....	741
Regulácia servopohonov Dynamixel pri pohybe robotického manipulátora .....	747
Využitie integrácie formálnych metód pri návrhu a analýze diskretných systémov .....	752
Secure Server Prototype based on Micro Kernel.....	759
Interactive 3D Visualization and Organization of Objects in Web Environment.....	764
Nový optický vláknový refraktometrický systém ovládaný cez Web.....	769
Rozhranie k inteligentnému audiovizuálnemu archívu .....	774
Algoritmus rozdelenia kódovej knihy v procese vektorového kvantovania s treliovým kódovaním.....	780
Výber kandidátskych uzlov na báze teórie hier v hybridných MANET - DTN.....	786
Využitie dôvernosti pri výbere kandidátskych uzlov v hybridnej MANET-DTN .....	792
PROPHET protocol for DTN in OPNET modeler .....	798
Riešenie elektromagnetických polí s využitím ANSYS .....	804
Hodnotenie ekonomickej efektívnosti bioplynovej elektrárne .....	810
Interactive Map of Campus for Android Operating System .....	816
Interactive ISO-OSI Protocol Model and its Utilizing in Education .....	820
Meranie elektrofyzikálnych vlastností izolácie olej-papier.....	825
Use of EMTP-ATP for Transients Calculation .....	830
Návrh a optimalizácia elektrických vykurovacích zariadení budov .....	836
Analýza teplotného poľa elektroenergetických zariadení .....	841
Nonlinear Effects in Optical Fiber and Wavelength Division Multiplexing .....	847

Biometric Systems And Face Recognition.....	853
Introduction To NoSQL Databases.....	858
Bagging Efficiency of Machine Learning in Steganalysis.....	864
Metódy oceňovania prenosových služieb .....	869
Analýza veľkých dát použitím technológie IBM InfoSphere BigInsights .....	875
Procesná analýza výrobnjej firmy s využitím vybraných technológií .....	880
Tvorba reportovacej vrstvy v prostredí Microsoft SQL Server 2012.....	886
Úlohy objavovania znalostí v prostredí technológie Microsoft SQL Server 2012.....	891
Analýza meteorologických dát s využitím nástroja QlikView .....	896
Analýza postojov v neformálnych krátkych textoch.....	901
Vyhľadavanie názorového spamu .....	906
Blocks of Cognitive Radio System .....	910
Návrh a implementácia paralelných programov s využitím JCSP .....	914
Processing of Measured Data with a Focus on Visibility .....	920
Experimentálny model PON siete pre areál TUKE .....	926
Kombinované banky filtrov.....	932
Konverzia časovej lifting implementácie diskkrétnej waveletovej transformácie do z-roviny .....	937
Vizualizačná platforma VizAlgo a jej rozšírenie .....	943
Co-existence of DVB-T/2 and LTE800 .....	948
Kognitívne rádio.....	954
Segmentácia obrazov metódou watershed.....	959
Implementácia metód segmentácie obrazov .....	964
Kalibrácia ohniskových vzdialeností kamery.....	969
Finančná predikcia vývoja vybraných menových párov Forex-u .....	974
Návrh a implementácia nástroja pre podporu predspracovania dát časových radov .....	978
Analýza možností výberu dát z databázy ORACLE .....	982
Návrh a realizácia riadenia modelu inkubátora pre podporu rastu húb pomocou jednočipového mikropočítača .....	987
Ohrev DPS vnorenými rezistormi.....	991

# Optimal State Control of the Mechatrical Laboratory Model B&P\_KYB

*Matej ORAVEC, Anna JADLOVSKÁ*

Department of Cybernetics and Artificial Intelligence, Faculty of Electrical  
Engineering and Informatics, Technical University of Košice, Slovak Republic

matej.oravec@tuke.sk, anna.jadlovaska@tuke.sk

**Abstract** — The article describes a new mechatrical laboratory model B&P\_KYB, which is specific in the construction and used construction parts. This article is focused on the modeling of the mechatrical laboratory model B&P\_KYB based on physical laws. Also it is focused on the creation, the analysis and the validation of the simulation model B&P\_KYB with the real laboratory model. For the trajectory tracking optimal state LQ control algorithm with an integrator in discrete form (LQI) has been implemented. This algorithm is based on the parameters of the state model of the mechatrical system B&P\_KYB. For the LQI algorithm has to be chosen a correct control scheme. The LQI control algorithm is verified on the real and the simulation mechatrical model B&P\_KYB. The article presents only a part of several results solved in thesis [5].

**Key words** — Ball and plate model, optimal LQ control, optimal LQ control with an integrator in discrete form

## I. INTRODUCTION

The mechatrical laboratory model B&P\_KYB is the model of the ball on the plate, it exists a lot of different variants of this kind of the mechatrical model (e. g. model of Humusoft CE151 in [1], [6], [7]). The model B&P\_KYB has a lot of specific components, which are unique. For the plate sloping are used servomotors from hobby models, but those have disadvantage, because their parameters aren't known and using of the experimental methods of the identification is necessary (in model of Humusoft CE151 are used stepmotors, in [6], [7]). The connection of the servomotors with the control PC is being done by using a single-chip microcomputer, what is different from model Humusoft CE151, which is using more expensive Lab card, in [6], [7]. The next difference is the use of the simple web camera (frames per second have to be more than 30). The advantage of the model B&P\_KYB is the possibility to use any web camera with similar parameters, but it is necessary to implement own algorithm for the image processing (in model of Humusoft CE151 are used functions of Matlab Image processing toolbox).

In fact, the model B&P\_KYB is different model than another Ball and plate model (e. g. Humusoft CE151 - [6], [7]), so one of the purpose of this article is to present to creation of the simulation model of B&P\_KYB, where the parameters of the simulation model of B&P\_KYB can be used for design and verification of the control algorithms based on input/output or state space model of B&P\_KYB, before they will be used on the real model. In this article is presented the LQI algorithm based on state space model of B&P\_KYB. Another types of the control algorithms of the mechatrical model B&P\_KYB are presented in [5].

## II. LABORATORY MODEL OF B&P\_KYB

### A. Description of the laboratory model B&P\_KYB

The laboratory mechatrical model B&P\_KYB (Fig. 1) is a real model, which is located in the Laboratory of mechatrical systems V142, Department of Cybernetics and Artificial Intelligence, FEEI, TU (<http://kyb.fei.tuke.sk/laben/miest/V142.php>). The model consists of two servo motors, plate, ball, web camera and single-chip microcomputer. The single-chip microcomputer is interface for the control of the servomotors from PC. The servomotors are rotating by the plate and ball is moving on the plate. The web camera is being fixed above the

plate and capturing the position of ball. For calculating of the ball position on the plate, an algorithm for the image processing in the language *C#* using libraries of the *emguCV*.

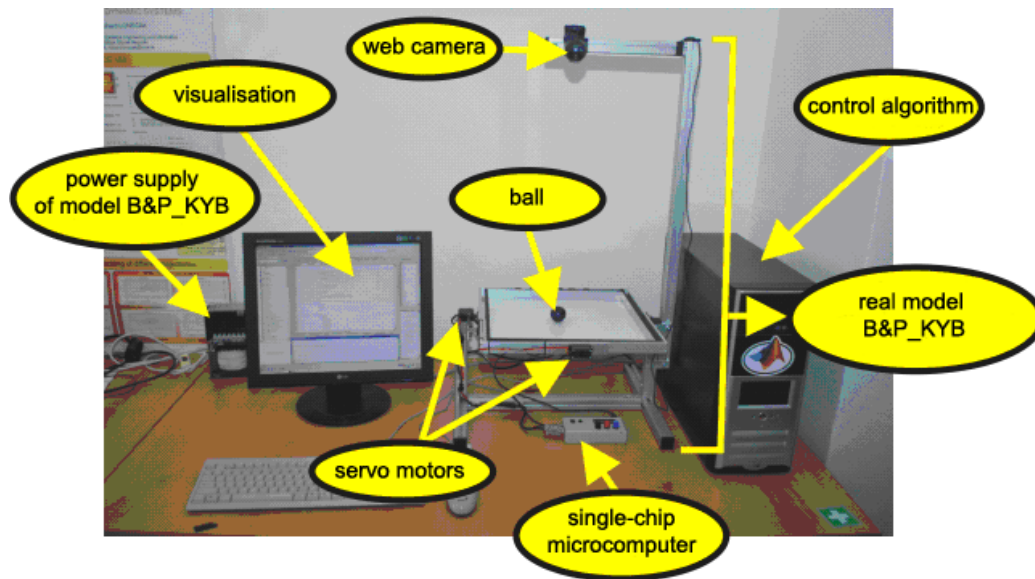


Fig. 1 Mechatronics laboratory model B&P\_KYB – Laboratory of mechatronics systems V142

The scheme of the laboratory model B&P\_KYB assumes, that the system is divided to two subsystems. First subsystem of the model consists of the servomotors, second subsystem consists of the ball on the plate, Fig. 2. The next assumption is that this laboratory model is not being taken as MIMO system, but as two SISO systems. Based on this fact, the system can be divided to two separate system, one for the direction of the axis  $x$  and second for the direction of the axis  $y$ .

The input of the subsystem *Servomotors* for the direction of the axis  $x/y$  is reference angle  $\alpha_x(t)/\beta_y(t)$ , output is an angle  $\alpha(t)/\beta(t)$ . An angle  $\alpha(t)/\beta(t)$  is input for the subsystem *B&P* for the direction of the axis  $x/y$ , output of this subsystem is  $x/y$  direction of the ball position  $y_x(t)/y_y(t)$ .

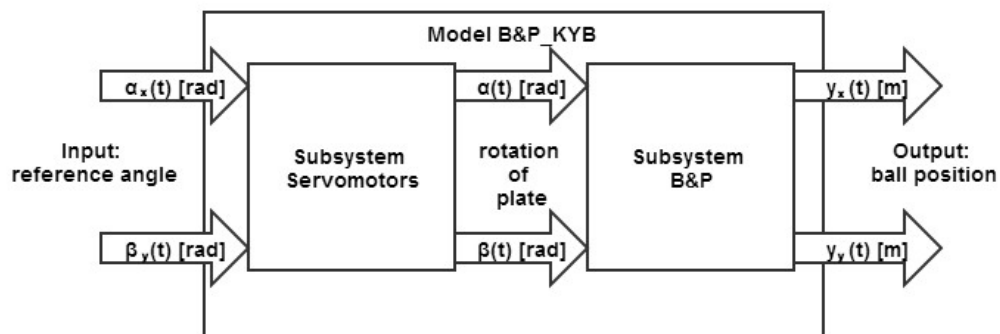


Fig. 2 Scheme of the mechatronics laboratory model B&P\_KYB

How has been described, the ball position  $y_x(t), y_y(t)$  is being captured by web camera above the plate. For using of the web camera it has been important to design and implement algorithm for the image processing (Fig. 3), which the output is the ball position  $y_x(t), y_y(t)$ . The servomotors are communicating with the PC with using of the single-chip microcomputer. The communication is realized by RS232 interface protocol. In another real mechatronics model, as for example Humusoft CE 151 in [6], [7], is used for communication the Lab card (not the single-chip microcomputer) with Matlab RealTime toolbox.



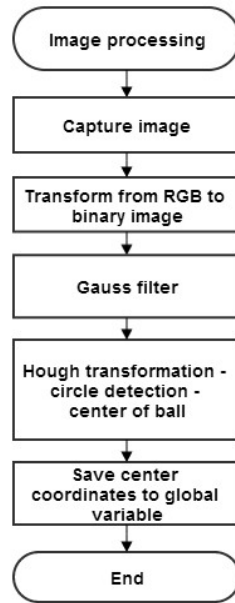


Fig. 3 Image processing algorithm of the mechatronical laboratory model B&P\_KYB

### B. Simulation model of B&P\_KYB

The Simulation model is based on the mathematical description of real model by differential equations with using the physical laws. The simulation model of B&P\_KYB should be used in situation, when the real model is not available or control algorithm could be break the real laboratory model.

Model B&P\_KYB consist of two subsystems, subsystem *Servomotors* and subsystem *B&P*. The subsystem *Servomotors* can be described by the linear differential equations:

$$\dot{\alpha}(t) = \frac{K_x}{T_{sx}} \cdot (\alpha_x(t) - \alpha(t)), \quad (1)$$

$$\dot{\beta}(t) = \frac{K_y}{T_{sy}} \cdot (\beta_y(t) - \beta(t)). \quad (2)$$

where  $K_x$ ,  $K_y$  are gain coefficients of the servomotors and  $T_{sx}$ ,  $T_{sy}$  are time constant of the servomotors. The parameters of the servomotors in the linear differential equations are obtained by identification from their step response by the method of surface [4].

The subsystem *B&P* describes two nonlinear differential equations. For their derivation are used Lagrange equations of the second kind. The parameters of the model B&P\_KYB are a ball radius  $r$  [m], a ball mass  $m$  [kg] and a plate length  $L$  [m]. The subsystem *B&P* describes equations in the direction of the axis  $x$  and  $y$ , more in [5]:

$$\left( \frac{J}{r^2} + m \right) \cdot \ddot{y}_x(t) - m \cdot g \cdot \sin \alpha(t) = 0 \quad \rightarrow \quad \ddot{y}_x(t) = \frac{5}{7} g \sin \alpha(t), \quad (3)$$

$$\left( \frac{J}{r^2} + m \right) \cdot \ddot{y}_y(t) - m \cdot g \cdot \sin \beta(t) = 0 \quad \rightarrow \quad \ddot{y}_y(t) = \frac{5}{7} g \sin \beta(t), \quad (4)$$

where for the moment of inertia:  $J = \frac{2}{5} mr^2$ .

The plate of the mechatronical model B&P\_KYB, where the ball is moving, has the form of square of dimensions  $\langle -L/2, L/2 \rangle$ . In the simulation model B&P\_KYB those limitations are being incorporated.

For verification of the simulation model with the real laboratory model was used the same step input signal. The verification is illustrated in Fig. 3 and Fig. 4, which shows, that response are nearly same and the simulation model is suitable replacement of the real model for the simulation and control algorithm testing. This simulation model is in the form of Grey box model.

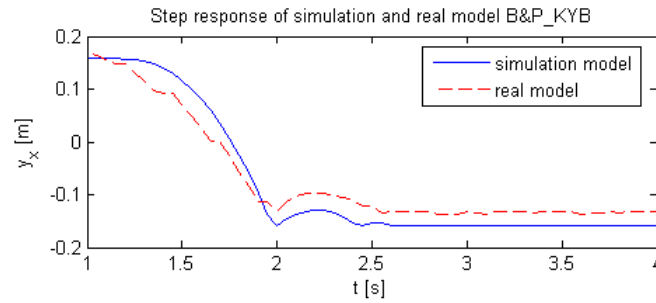


Fig. 3 Comparison of the simulation and the real model – direction of the axis x

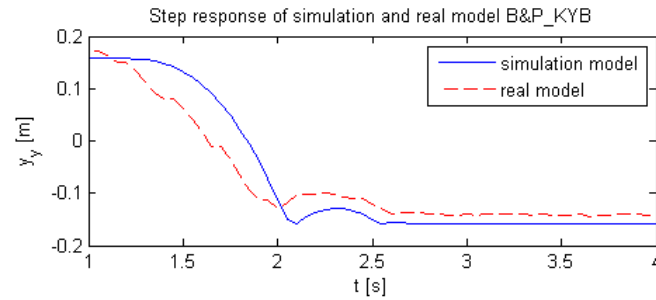


Fig. 4 Comparison of the simulation and the real model – direction of the axis y

### III. THE OPTIMAL STATE LQ CONTROL OF THE REAL AND THE SIMULATION MODEL B&P\_KYB

For the control B&P\_KYB laboratory model it is possible to use many different control algorithms based on input/output or state space model of B&P\_KYB and design of some control algorithm and their verification on simulation and real model are shown in [5]. One of the control algorithm, which can be used for simulation or real model B&P\_KYB, is optimal state LQ control algorithm. This type of the optimal control is used for feedback control if the goal of the control is an equilibrium state. If the goal of the control is the reference trajectory tracking, it is necessary to implement the algorithm of LQI. For design of LQI algorithm is important to define state space description of the model B&P\_KYB. The control scheme of the LQ and the LQI control is different. The control scheme of the LQI control is shown on Fig. 6, more is presented in [5].

#### A. State space model B&P\_KYB

For the creation of state space model it is important to define state variables. The vector of the state variables for direction of the axis x is:

$$\bar{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} y_x(t) \\ \dot{y}_x(t) \\ \alpha(t) \end{bmatrix} = \begin{bmatrix} \text{position [m]} \\ \text{velocity [m} \cdot \text{s}^{-1}] \\ \text{angle [rad]} \end{bmatrix} \quad (5)$$

Input vector has the form:

$$u(t) = [\alpha_x(t)] \quad (6)$$

From the equation (1) and (3) is able to define substitution in canonical form (direction of axis x):

$$y_x(t) = x_1(t)$$

$$f_1: \dot{x}_1(t) = x_2(t) \quad (7)$$

$$f_2: \dot{x}_2(t) = \frac{5}{7} g \sin x_3(t) \quad (8)$$

$$f_3: \dot{x}_3(t) = \frac{K_x}{T_{sx}} \cdot (\alpha_x(t) - x_3(t)) \quad (9)$$

For linearization of the system B&P\_KYB in direction of the axis x the reference angle is  $\alpha_x(t) = 0$ , the ball position  $y_x(t) = 0$  and the velocity  $\dot{y}_x(t) = 0$ . The linearization of the system is realized in the equilibrium point, which values are  $x_{xEP} = [x_1 = 0 \quad x_2 = 0 \quad x_3 = 0]$ .

Continuous state space of the model B&P\_KYB (direction of the axis x):

$$\begin{aligned} \dot{x}_x(t) &= A_x x_x(t) + B_x u_x(t) \\ y_x(t) &= C_x^T x_x(t) \end{aligned} \quad (10)$$

The matrix of the system dynamic could be defined by the Jacoby matrix:

$$A_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{bmatrix}_{x_{XEP}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & k_g g \\ 0 & 0 & -\frac{K_x}{T_{sx}} \end{bmatrix}_{x_{XPB}}, \quad k_g = \frac{5}{7} \quad (11)$$

The input matrix is:

$$B_x = \begin{bmatrix} \frac{\partial f_1}{\partial u} \\ \frac{\partial f_2}{\partial u} \\ \frac{\partial f_3}{\partial u} \end{bmatrix}_{x_{XPB}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{x_{XPB}} \quad (12)$$

The process to define state space of the system for direction of the axis  $y$  is the same like for the axis  $x$ . The output matrices for both directions are:  $C_x^T = C_y^T = [1 \ 0 \ 0]$ .

For the mechatronical model B&P\_KYB the LQI algorithm will be designed in the discrete form. Sample period will be  $T_s = 0,05 \text{ s}$ . State space equations of the model in the discrete form are following:

$$\begin{aligned} x_x(i+1) &= F_x x_x(i) + G_x u_x(i) \\ y_x(i) &= C_x^T x_x(i) \end{aligned}, \quad (13)$$

$$\begin{aligned} x_y(i+1) &= F_y x_y(i) + G_y u_y(i) \\ y_y(i) &= C_y^T x_y(i) \end{aligned}, \quad (14)$$

where system matrices  $F_x, F_y$  are (3x3) and input matrices  $G_x, G_y$  are (3x1). Output matrices are same like in the continuous time state space matrices.

### B. Design and implementation of the LQI algorithm

Design of the optimal LQ algorithm assumed, that quadratic functional has been minimized:

$$J_{LQ} = x^T(M) Q x(M) + \sum_{i=0}^{M-1} (x^T(i) Q x(i) + x^T(i) S u(i) + u^T(i) S^T x(i) + u^T(i) R u(i)) \quad (15)$$

where  $M$  is finite horizon. Functional  $J_{LQ}$  has been minimized with respect to the limitations, which actually are the discrete state equations of the system. Matrix  $Q$  of the functional is the weight matrix of the state variables (positive semidefinite),  $R$  is the input weight matrix (positive definite) and  $S$  is the zero matrix (dimension (3x1)). The values of these matrices are optional, but have to suit those mentioned conditions. To determine the state feedback gain  $k_x$  it is necessary to solve the discrete Riccati algebraic equation [2]:

$$P(i) = Q + F^T P(i) F - [S + F^T P(i) G] [R + G^T P(i) G]^{-1} [S + F^T P(i) G]^T, \quad (16)$$

and the relation for the state feedback gain is:

$$k_x(i) = [R + G^T P(i) G]^{-1} [S + F P(i) G]^T. \quad (17)$$

This solution of the state feedback amplification could be used only for the goal of the control to equilibrium.

For reference trajectory tracking it's necessary to add discrete form of an integrator of control deviation  $S_u$ . The detail solution for an integrator in the discrete form could be find in [3], for the calculation of an integrator in discrete form in step  $i+1$  applies:

$$\begin{aligned} S_u(i+1) &= S_u(i) + K_2 e(i) = S_u(i) + K_2 (w(i) - y(i)) = \\ &= S_u(i) + K_2 (w(i) - Cx(i)) \end{aligned}, \quad (18)$$

for new control input:

$$u(i+1) = -K_1 x(i+1) + S_u(i+1) + K_2 w(i), \quad (19)$$

where  $w(i)$  is reference trajectory and its value in step  $i$  and  $K_1, K_2$  are the matrices calculated from this relation:

$$[-K_1 \quad -K_2] = [-k_x F \quad -I - k_x G] \begin{bmatrix} F - I & G \\ C & 0 \end{bmatrix}^{-1}, \quad (20)$$

where the feedback gain  $k_x$  represent steady value of the solving (17).

Base on these steps the control algorithm for the simulation and the real model of the mechatronical system B&P\_KYB is being designed. This algorithm is being shown in the flowchart in Fig. 5. This control algorithm is implemented in control scheme in Fig. 6.

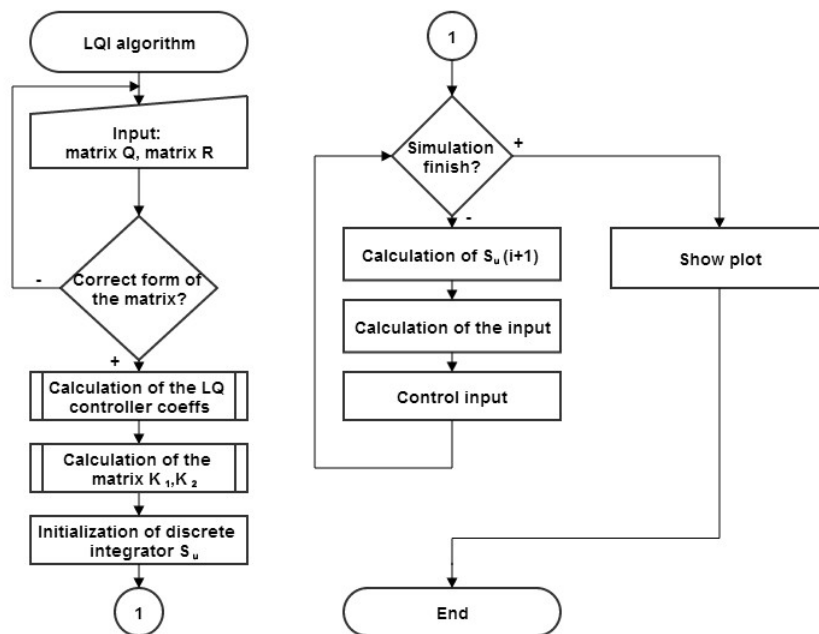


Fig. 5 Flowchart of the LQI algorithm

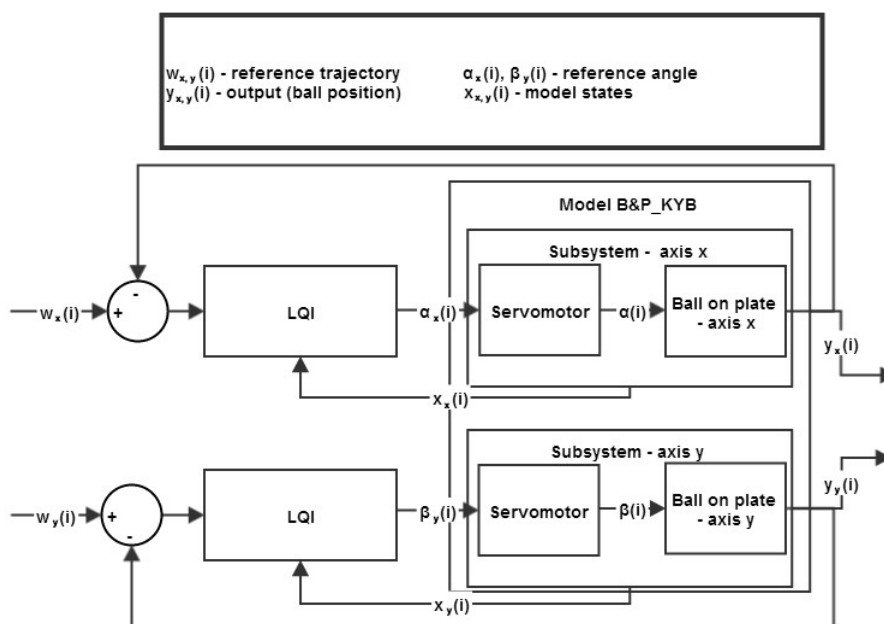


Fig. 6 Control scheme for the reference trajectory tracking with using the LQI algorithm

### C. Verification of the implemented LQI algorithm

The LQI algorithm is implemented in the form of newly programmed function *paramLqsC* in *Matlab*. The simulation results of the LQI algorithm for the reference trajectory tracking (the circle, the square) are shown in the common Fig. 7, Fig. 8.

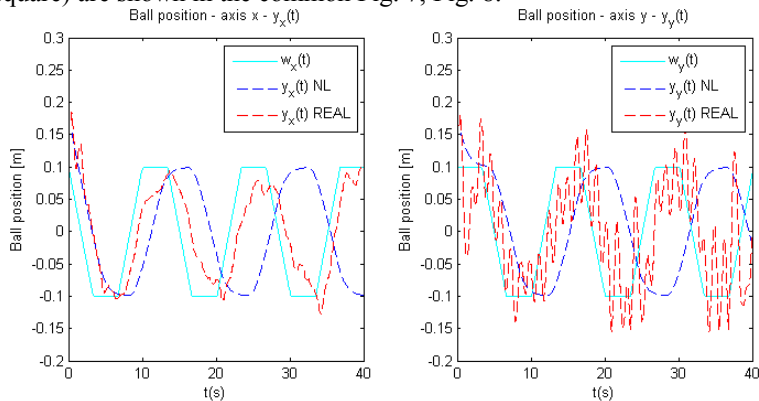


Fig. 7 Comparison of the simulation and the real model – the square trajectory tracking

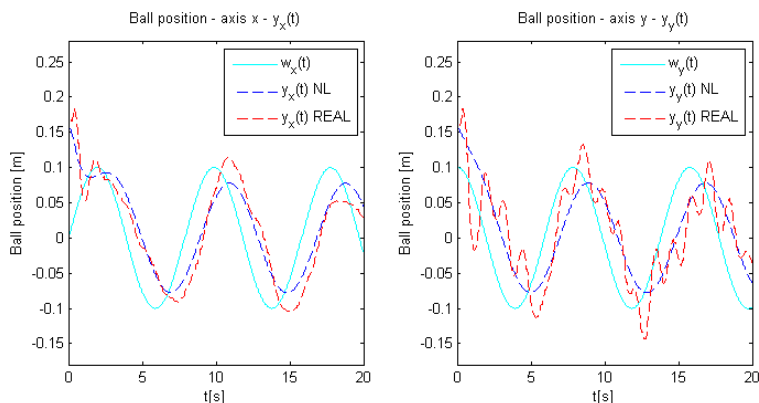


Fig. 8 Comparison of the simulation and the real model – the circle trajectory tracking

## IV. CONCLUSION

This article presents the laboratory model B&P\_KYB in the terms of the structural design, the process of the creating simulation model B&P\_KYB based on the physical laws and design of LQI control. The simulation model is created in the form Grey box model, which is one of many forms of creating the simulation model (in [5] the Black box model form is being presented too). This simulation model after verification with the real model is suitable replacement of the real model and it is possible to use for simulation and verification of the control algorithms. The state-space model at the selected equilibrium point served for the design of LQI algorithm of the model B&P\_KYB for the tracking of chosen reference trajectory. The simulation results of LQI algorithm form are tolerable, but response of real or simulation model B&P\_KYB on the reference trajectory changing is with the delay.

Communication between the real model and the PC (data acquisition and exchange) with the usage of the web camera and the single-chip microcomputer is very effective. The model using basic, but effective components enables adaptation of the model for actual conditions and requirements. The other models are more complicated, what is limitation for user.

## ACKNOWLEDGEMENTS

This contribution is the result of the Vega project implementation – Dynamic Hybrid Architectures of the Multi-agent Network Control Systems (No. 1/0286/11), supported by the Scientific Grant Agency of Slovak Republic – 40%, and of the KEGA project implementation – CyberLabTrainSystem – Demonstrator and Trainer of Information-Control Systems (No. 021TUKE-4/2012) – 30%, and of the Research and Development Operational Program for project: University Science Park Technicom for innovative applications with knowledge technology support, (ITMS code 26220220182), co-financed by the ERDF – 30%.

REFERENCES

- [1] A. Jadlovská. – Š. Jajčišin – R. Lonščák, “Modelling and PID Control Design of Nonlinear Educational Model Ball & Plate,” in *Proceedings of the 17th International Conference on Process Control '09*, Štrbské Pleso, High Tatras, Slovakia, 2009, pp. 475–483, ISBN 978-80-227-3081-5. web: <<http://www.kirp.chtf.stuba.sk/pc09/data/abstracts/038.html>>
- [2] D. Krokavec – A. Filasová, *Discrete systems*, Košice, ELFA, 2006. ISBN 80-8086-028-9
- [3] O. Modrlák, *Basics of analysis and synthesis in state space*. Materials for study, Liberec: TU – FM, 2004.
- [4] P. Noskievič, *Modeling and identification of systems*, Ostrava, Montanex, 2007. ISBN 80-7225-030-2
- [5] M. Oravec, *Control of Mechatronical Systems using .NET Applications*, Master thesis, Supervisor: doc. Ing. Anna Jadlovská, PhD. Košice: TU – FEI, 2014.
- [6] <http://kyb.fe.i.tuke.sk/laboratoria/modely/gnp.php>
- [7] <http://humusoft.cz/produkty/models/>

# Principles of polymorphic types

<sup>1</sup>Marián POLÁK, <sup>2</sup>Peter PRAZŇÁK, <sup>3</sup>Valerie NOVITZKÁ

<sup>1</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>2</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>3</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>marian.polak@student.tuke.sk, <sup>2</sup>praznak@minv.sk, <sup>3</sup>valerie.novitzka@tuke.sk

**Abstract** – Our paper deals with polymorphic types that are often used in various programming paradigms. We shortly introduce their syntax, typing rules and semantics. The main aim of our work is to construct a categorical model of the System F for educational purposes because categorical model can appropriately illustrate principles of polymorphic types in graphical style. Our work can contribute not only to deeper knowledge of modeling polymorphic types but our results will be used in future research concerning with modeling component based program systems, where polymorphic types are often used.

**Keywords** – category theory, polymorphic types, System F

## I. INTRODUCTION

In general, the concept *polymorphism* is defined as ability to gain more than one form and it means something like a "variety" or "manifoldness". Polymorphism is well known and often used in various areas of informatics. There are several kinds of polymorphism, e.g. ad-hoc polymorphism, inclusion polymorphism (subtyping), intensional polymorphism and many others [1]. In this paper we concern in parametric polymorphism, the mostly used kind of polymorphism. We use the definition of System F introduced by Girard in [2].

The article is divided into two parts. The first part contains definitions of basic notions of type theory needed for understanding our approach. In the second part we construct categorical model of polymorphic types as a lifting over basic category of kinds. The category of polymorphic types is then a category of small categories, where objects are small categories of types and morphisms are substitution functors. We illustrate our results in figures for better understanding.

## II. BASIC NOTIONS

Every program can be considered as a collection of data structures and algorithms working with these structures. Every data structure has assigned a type. A *type* is a collection of values that can be represented in a computer together with a set of operations on it. A *type system* is a part of every modern programming language [3]. Type checking is a necessary part of semantical phase of compilers.

Type systems are obviously formulated in  $\lambda$ -calculus introduced by Alonzo Church in the first part of 20th century. It was extended to simply typed  $\lambda$ -calculus and later enriched by other types as reference types, subtypes, recursive types and polymorphic types.

Syntax of polymorphic types is defined by the following production rules:

$$\begin{aligned} t &::= x \mid \lambda x : T.t \mid t t \mid \lambda X.t \mid t[T] \\ v &::= \lambda x : T.t \mid \lambda X.t \\ T &::= X \mid T \rightarrow T \mid \forall X.T \\ \Gamma &::= \emptyset \mid \Gamma, x : T \mid \Gamma, X \end{aligned}$$

The metavariable  $t$  denotes terms that can be variable  $x$ ,  $\lambda$ -abstraction (function definition)  $\lambda x : T.t$  and application  $t t$ . The last two forms belong to polymorphic types and they define type abstraction  $\lambda X.t$  and type application  $t[T]$ . Type abstraction enables to define function where a type of parameter is type variable  $X$  substitutable by appropriate type and type application expresses an instantiation of term type, i.e. substitution of a type  $T$  on a function expressed by a term  $t$ .

The metavariable  $v$  denotes values and the metavariable  $T$  denotes types. In the production rule for types are defined possible forms of type expressions.

The last production rule defines possible forms of type contexts  $\Gamma$  containing sequence of typed variables declarations and type variables occurring in a term. A term has a form of sequent

$$\Gamma \vdash t : T$$

Type checking uses typing rules of the following form:

$$\frac{x : T \in \Gamma}{\Gamma \vdash x : T} \text{ (T-var)}$$

$$\frac{\Gamma, x : T_1 \vdash t_2 : T_2}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2} \text{ (T-abs)}$$

$$\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \quad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1 t_2 : T_{12}} \text{ (T-app)}$$

$$\frac{\Gamma, X \vdash t_2 : T_2}{\Gamma \vdash \lambda X. t_2 : \forall X. T_2} \text{ (T-Tabs)}$$

$$\frac{\Gamma \vdash t_1 : \forall X. T_{12}}{\Gamma \vdash t_1[T_2] : [X \mapsto T_2]T_{12}} \text{ (T-Tapp)}$$

The first three rules are known typing rules for typed  $\lambda$ -calculus. The last two rules serve for polymorphic types.

The semantics of polymorphic types is defined by evaluation rules describing term reduction. A step of term reduction is expressed by the symbol  $\rightarrow$ . Evaluation rules are the following:

$$\frac{t_1 \rightarrow t'_1}{t_1 t_2 \rightarrow t'_1 t_2} \text{ (E-app1)}$$

$$\frac{t_2 \rightarrow t'_2}{v_1 t_2 \rightarrow v_1 t'_2} \text{ (E-app2)}$$

$$(\lambda x : T_{11}. t_{12}) v_2 \rightarrow [x \mapsto v_2] t_{12} \text{ (E-appabs)}$$

$$\frac{t_1 \rightarrow t'_1}{t_1[T_2] \rightarrow t'_1[T_2]} \text{ (E-T-app)}$$

$$(\lambda X. t_1)[T_2] \rightarrow [X \mapsto T_2]t_1 \text{ (E-T-appT-abs)}$$

The first three rules are obvious evaluation rules for typed  $\lambda$ -calculus. The fourth rule expresses that if a term  $t_1$  is reducible to a term  $t'_1$ , the instantiation with the type  $T_2$  is reducible in the same manner. The second rule defines how a type variable  $X$  is substituted by a type  $T_2$ .

### III. CATEGORICAL MODEL OF POLYMORPHIC TYPES

In this section we explain in detail how we construct our model of polymorphic types. In the first subsection we formulate polymorphic signature, then we construct classifying category of kinds and then we define lifting functor for modeling polymorphic types.

#### A. Polymorphic signature

To formulate polymorphic signature we introduce the new notion of type kind. *Type kind*  $K$  is a set of all types substitutable for a type variable. In other words, a type variable  $X : K$  can be substituted only with some type  $T$  from  $K$ ,  $T \in K$ . Kinds  $K_1$  and  $K_2$  can contain the same types, these sets need not be disjoint. All kinds form the *universal type*  $\mathcal{T}ype = \forall X. T$ . The notation

$$X : K$$

associates a kind  $K$  to type variable  $X$ . Now we can define polymorphic signature.

A *polymorphic signature* is ordered pair

$$(\Sigma, (\Sigma_k))$$



where

- $\Sigma$  is a generic signature;
- $\Sigma_k$  is a type signature for finite sequence of kinds  $k = K_1, \dots, K_n$ .

A *generic signature*  $\Sigma$  [4] is an ordered pair

$$\Sigma = (K, F)$$

where

- $K$  is a finite set of kinds,
- $F$  is finite set of functional symbols of the form  $F : K_1, \dots, K_n \rightarrow L$  between kinds.

For every finite sequence of kinds  $k = K_1, \dots, K_n$  we define a *type signature* as an ordered pair:

$$\Sigma_k = (Term, OP)$$

where

- $Term$  is a finite set of types  $T$

$$X_1 : K_1, \dots, X_n : K_n \vdash T : Type$$

constructed on the basis of  $\Sigma$ ;

- $OP$  is a finite set of operations  $f : T_1, \dots, T_n \rightarrow T$ .

We introduce also a *kind context*  $\Delta$ :

$$\Delta = X_1 : K_1, \dots, X_n : K_n$$

as a sequence of declarations of type variables according to kinds.

We note that now we have two contexts [5]

- kind context with type variables of particular kinds introduced above and
- type context defined in the previous section

$$\Gamma - x_1 : T_1, \dots, x_m : T_m.$$

Now, we have a new form of a term in polymorphic type theory as a sequent

$$\Delta, \Gamma \vdash t : T$$

with two contexts.

### B. Classifying category of kinds

When we have defined kinds of types, we can construct classifying category of kinds  $Cl(\Sigma)$  over generic signature  $\Sigma$ . *Category of kinds*  $Cl(\Sigma)$  consists of:

- objects are kind contexts  $\Delta_1, \Delta_2, \dots, \Delta_l$  considered as finite products of kinds

$$K_1 \times K_2 \times \dots \times K_l$$

- morphisms are sequences of terms  $t_1, \dots, t_n$ , for  $i = 1, \dots, n$  such that every term  $t_i$  has the form

$$\Delta_1 \vdash t_i : K_i$$

where  $K_i$  is a kind from  $\Delta_2$ .

Category of kinds is in Fig.1.

We have proved that the category  $Cl(\Sigma)$  of kinds constructed above fulfils the axioms of category, i.e. every object has identity morphism, composition of morphisms is associative and the set of morphisms between two kinds form homset. It is clear that  $Cl(\Sigma)$  has to be a category with finite products.

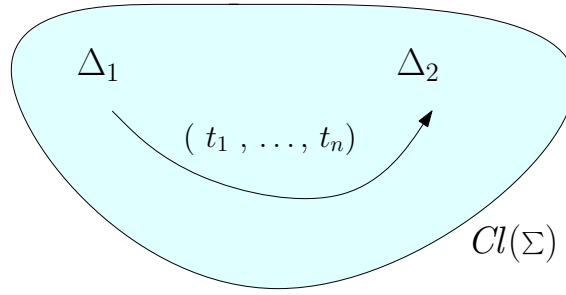


Fig. 1 Category of kinds

### C. Model of polymorphic types

The category  $Cl(\Sigma)$  still does not include all needed information about possible substitutions of kinds by appropriate types. Therefore, we construct the category of polymorphic types  $Cat$  above the category of kinds  $Cl(\Sigma)$  by means of functor  $Lift$  in the following way. The functor  $Lift$  we specify as

$$Lift : Cl(\Sigma)^{op} \rightarrow Cat$$

which means that functor of lifting has as a domain dual category of kinds and as a codomain the category  $Cat$  of small categories. Instead of approach published in [6] we define the functor  $Lift$  for objects and morphisms of category of kinds  $Cl(\Sigma)$  in simpler manner.

Functor  $Lift$  in  $Cl(\Sigma)$  assigns

- to every object  $\Delta$  a category of types  $C_\Delta$ .

Let  $\Delta$  be a kind context, an object in  $Cl(\Sigma)$ . We construct the corresponding category  $C_\Delta$  as follows:

- objects are types  $T$  in kind context  $\Delta$ , what can be written as

$$\Delta \vdash T : Type$$

- morphisms are terms of the form

$$\Delta, \Gamma \vdash t : T$$

of corresponding types in  $\Delta$ .

For every kind context  $\Delta$  we construct corresponding category of types  $C_\Delta$ . Concrete accessible types are assigned to particular kind context within this category of types. Again, we proved that such categories of types satisfy axioms of categories. A category of types is in Fig.2

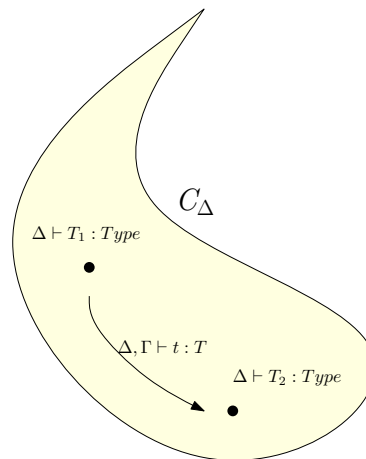


Fig. 2 Category of types for kind context  $\Delta$

Now we can define the functor  $Lift$  also for morphisms from classifying category of kinds to morphisms between type categories in  $Cat$ .

Let

$$\begin{aligned}\Delta_1 &= (X_1 : K_1, \dots, X_n : K_n) \text{ and} \\ \Delta_2 &= (Y_1 : L_1, \dots, Y_m : L_m)\end{aligned}$$

be kinds, i.e. objects in the category  $Cl(\Sigma)$ . Let  $T_1, \dots, T_m$  be types in kinds  $L_1, \dots, L_m$ , respectively, i.e. for all  $i = 1, \dots, m$  we can write

$$\Delta_1 \vdash T_i : L_i.$$

Functor *Lift* assigns to every morphism

$$\Delta_1 \rightarrow \Delta_2$$

a functors *Subst* between corresponding type categories in *Cat*:

$$Subst : C_{\Delta_2} \rightarrow C_{\Delta_1}.$$

by usage of type substitution as follows. For every type  $T$  in  $\Delta_2$ :

$$\Delta_2 \vdash T : Type$$

we assign a type in which all occurrences of type variables  $Y_1, \dots, Y_m$  are substituted by the types

$$T_1 \in K_1, \dots, T_n \in K_n$$

so that

$$\Delta_2 \vdash T : Type \mapsto [T_1/Y_1, \dots, T_m/Y_m]\Delta_1 \vdash T : Type.$$

Categorical model of polymorphic types is in Fig.3.

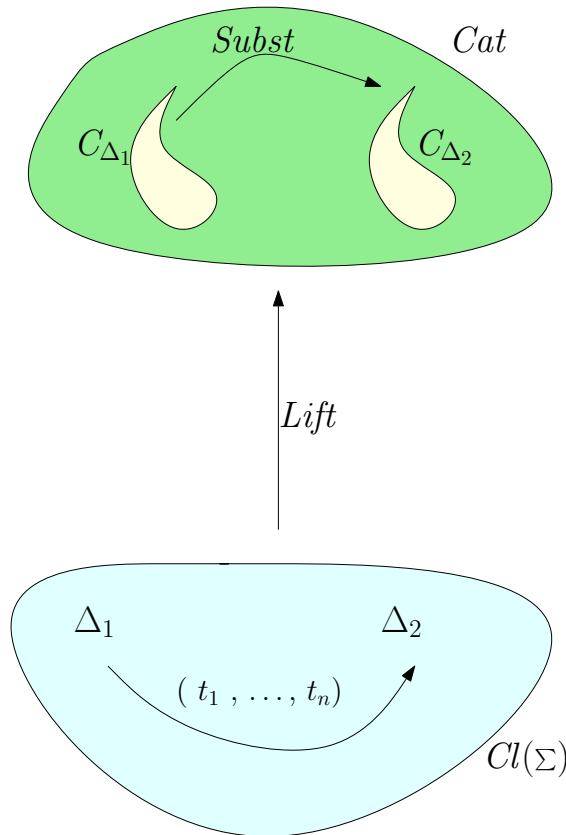


Fig. 3 Categorical model of polymorphic types

We note that opposite direction of substitution in category *Cat* substantiates usage of dual category  $Cl(\Sigma)^{op}$  in the definition of the functor *Lift* as it is used in [7].

#### IV. CONCLUSION

In this paper we present our approach to modeling polymorphic types by categories. Polymorphic types are often used in many programming languages, especially in object oriented paradigm. Therefore exact model that provides our paper can help not only to understand but also to examine important properties of polymorphic types in rigorous provable manner. We plan to use this model in our future research concerning with modeling component based program systems in the terms of categories.

#### ACKNOWLEDGMENT

This research was supported by the project APVV-0008-10: Modelling, simulation and implementation of GPGPU-enabled architectures of high-throughput network security tools

#### REFERENCES

- [1] J. COLLINS, *A history of the theory of types*. Germany: LAP Lambert Academic Publishing, June 2012.
- [2] J.-Y. Girard, Y. Lafont, and P. Taylor, *Proofs and Types*, ser. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1989, vol. 7.
- [3] B. PIERCE, *Types and programming languages*. Cambridge: MIT Press, Oct. 2002.
- [4] B. JACOBS, *Categorical logic and type theory*. Netherland: Elsevier Science, March 1999.
- [5] V. NOVITZKÁ and D. MIHÁLYI, "Polymorphic type theory as a base for categorical logic," *Acta Electrotechnica et Informatica*, vol. 7, pp. 19–25, September 2007.
- [6] V. NOVITZKÁ, D. MIHÁLYI, and V. SLODIČÁK, "Categorical Type Analysis for Parsing Algebras," *SAMI 2007*, vol. 7, pp. 477–485, September 2007.
- [7] V. NOVITZKÁ and V. SLODIČÁK, *Categorical structures and their applications in informatics*. Košice, Slovakia: Equilibria, Sept. 2010.

# Knižnica funkcií pre testovanie AD prevodníkov v LabVIEW

*Vojtech VITKOVIČ, Ján ŠALIGA*

Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical university of Košice, Slovak Republic

vojtech.vitkovic@student.tuke.sk, jan.saliga@tuke.sk

**Abstract** — Testovanie AD prevodníkov vyžaduje pomerne zložité matematické spracovanie dát získaným z testovania. Pre zjednodušenie vytvárania testovacích programov vyvinutá knižnica funkcií pre LabVIEW, ktorá obsahuje všetky potrebné postupy spracovania a vyhodnotenia výsledkov testovania v súlade so štandardom IEEE 1241-2010 a projektom DYNAD. Ako príklad praktického využitia vytvorenej knižnice bol zrealizovaný softvér v LabVIEW, ktorý okrem základným testovacích metód prezentuje aj rozdeiel vo výsledkoch testovania pri použití postupov doporučovaných štandardom IEEE a DYNAD. Tento softvér je dostupný aj prostredníctvom internetu na stránke Laboratória merania KEMT FEI TU v Košiciach.

**Keywords** — AD prevodník, statická metóda, histogram, nelinearita, dynamické parametre, SNR, SINAD, koherentné vzorkovanie, šum, spektrum, oknová funkcia, LabVIEW

## I. ÚVOD

Testovanie AD prevodníkov (ADP) predstavuje komplexnú a náročnú úlohu tak z hľadiska hardvérového riešenia testovacieho pracoviska a extrémnych požiadaviek na presnosť a kvalitu meracích prístrojov ako aj z hľadiska spracovania dát získaných z testovania. Práve pre zjednodušenie tohoto druhého kroku pri testovaní bola vytvorená a knižnica funkcií vychádzajúca zo štandardu IEEE 1241-2010 a výsledkov projektu DYNAD.

Niektoré štandardizované metódy sa odlišujú v jednotlivých odporúčaníach. Snahou vyselektovať vhodné riešenie odhaľujúce prenosové charakteristiky AD prevodníkov, bola analýza aplikovaná na dve odporúčané metódy testovania a to na americký štandard IEEE 1241-2010 a európsky projekt DYNAD.

Tento článok je organizovaný nasledovne: v prvej časti je poukázané, v ktorých častiach sa odlišujú spomínané odporúčania pre testovanie ADP. Jadro článku obsahuje samotné knižničné funkcie ako aj riešenie demonštračného programu. Navrhnutý vyhodnocovací softvér obsahuje virtuálny AD prevodník so simuláciou zdroja presného harmonického priebehu. Do programu boli integrované základné meracie metódy testovania. Finálna časť predstavuje aplikácia softvéru cez vzdialené pripojenie.

## II. DIFERENCIA ŠTANDARDOV

### A. Základné metódy testovania

Pri porovnávaní rôznych odporúčaní testovania sa jednotlivé metódy pre výpočet parametrov odhaľujúcich nelinearity a správanie sa AD prevodníkov, v niektorých črtách líšia, aj keď podstata výpočtov zostáva v princípe rovnaká.

Odlíšne návrhy pre výpočet je možné spozorovať už pri určovaní počiatočných podmienok pre korektný test histogramovej metódy. Napríklad pri výpočte dynamických parametrov v časovej oblasti štandard IEEE definuje vzorec pre efektívny počet bitov prevodníka (ENOB) presnejšie ako projekt DYNAD. Získavanie dynamických parametrov vo frekvenčnej oblasti je podľa projektu DYNAD zložitejšie. Výpočty sa odlišujú nielen pre koherentné, ale aj pre nekoherentné vzorkovanie, pričom európsky projekt definuje aj výpočet pre šumový prah (NFL). Pri nekoherentnom vzorkovaní sú jednotlivé zložky spektra po aplikovaní oknovej funkcie korigované ziskom danej oknovej funkcie ako je to určené americkým štandardom. DYNAD však túto korekciu prevádza tzv. opravným faktorom vypočítaným z parametrov použitej oknovej funkcie.

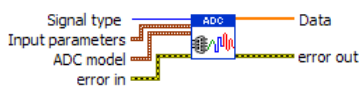
Spomínané rozdiely medzi štandardom IEEE 1241-2010 [1] a projektom DYNAD [2] sú v diplomovej práci [3] podrobnejšie analyzované.

### III. REALIZÁCIA KNÍŽNIČNÝCH FUNKCIÍ

Pri tvorbe knižničných funkcií bolo ťažisko položené najmä na komplexnosť funkcií a jednoduchosť ich využitia pri tvorbe finálneho programu. Vychádzalo sa pri tom už z existujúcich funkcií vytváraných ale nesystematicky pri experimentoch s ADP na KEMT FEI TU v Košiciach ako aj v rámci medzinárodného projektu s univerzitou SANNIO v Benevente v Taliansku. Tieto bolo potrebné zjednotiť a doplniť o chýbajúce časti, resp. opraviť stávajúce chyby v kódach.

Testovacie pracovisko bolo zrealizované s využitím dostupných knižníc matematických nástrojov v programe LabVIEW [4]. Získané programové knižnice bolo potrebné dôkladne štruktúrované analyzovať. Metódy testovania AD prevodníkov, ktoré boli naprogramované v jednotlivých častiach knižnice, bolo nutné porovnávať a overovať ich správnosť riešení. Cieľom bolo aj prerozdeliť jednotlivé odporúčania testovania a tým umožniť používateľovi programu si určiť, podľa ktorého štandardu, či projektu má test prebehnúť a vyhodnotiť výsledky.

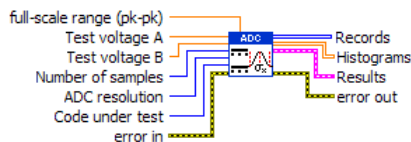
#### A. Simulácia hardvéru



Obr.1 Hardware simulation

Táto knižničná funkcia umožňuje vygenerovať dáta zo simulovaného testovania ADP. Okrem ideálneho ADP môže užívateľ zvoliť neideálny ADP v niektorej z troch typov architektúr so zvolenými hodnotami nelinearity a aditívneho vstupného šumu pri ľubovoľnom rozlíšení ADP. Táto funkcia teda v prípade potreby do istej miery nahrádza skutočný AD prevodník s architektúrou Flash, Successive approximation alebo Sigma-delta.

#### B. Statické metódy testovania



Obr.2 Express static test

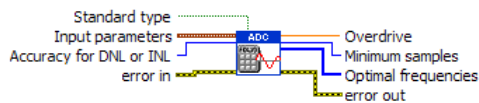
Táto komplexná funkcia v sebe zahŕňa statický test prevodníkov pomocou jednosmerných napätí pre analýzu jednotlivých kvantizačných úrovní testovaného prevodníka. Pomocou dvoch známych (odmeraných) jednosmerných napätí privedených na vstupy knižnice A a B, a ďalších potrebných parametrov, ako je uvedené na obr. 2, je možné získať hodnoty konkrétnych prechodových kvantizačných úrovní na základe podmienenej pravdepodobnosti výskytu jednotlivých kódov pre zvolené vstupné jednosmerné napätia. Výsledkom je aj zobrazenie histogramu kódov vo vstupnom rozsahu prevodníka. Táto funkcia bola doplnená o vstupný parameter full-scale range, pre možnosť určenia vstupného rozsahu podľa požiadaviek používateľa (v jednotkách Volt).

#### C. Histogramové metódy

Histogramové metódy testovania umožňujú určenie parametrov testovaného ADP pri jeho buzení známym signálom, najčastejšie harmonickým, analýzou histogramu výstupných kódov ADP spracovaných do formy histogramu. Samotné zobrazenie histogramu nám udáva počet vzoriek signálu pripadajúci na jednotlivý kód, v rámci celkového rozsahu prevodníka. Vypočítaná integrálna nelinearita (INL) indikuje odklon reálnej prevodovej charakteristiky od ideálnej priamky. Diferenciálna nelinearita (DNL) určuje odchýlku šírky kvantizačného kroku reálneho prevodníka od ideálnej šírky. Nelinearity INL a DNL je nutné rátať po prevedení korekcie zisku, pre INL je navyše potrebná korekcia offsetu [6].

Pre dostatočne presné merania je potrebné získať dostatočný počet vzoriek – štatistiku. Funkcia „Express test condition“ preto v sebe zahŕňa aj výpočet počiatočných podmienok pre korektný test nelinearít INL a DNL.

#### Výpočet počiatočných podmienok



Obr.3 Express test conditions

Pre vyhodnotenie histogramového testu bola navrhnutá funkcia umožňujúca výpočet z kombinácie viacerých záznamov  $R$ . Používateľ funkcie si aj jednoducho určí konfidenčnú hladinu  $B$  pre zvolenú presnosť určenia chyby kódového binu. Výpočet kritickej hodnoty  $Z_{N,u}$  je už realizovaný v programe. Výpočet optimálnych frekvencií je realizovaný na základe výpočtu počtu vzoriek v zázname  $M$  a vzorkovacej frekvencie  $f_s$ , hľadaním takých frekvencií, ktoré nemajú spoločného deliteľa s  $f_s$  (2) [5], aby bolo dosiahnuté koherentné vzorkovanie, pričom musí byť splnená Nyquistova podmienka.

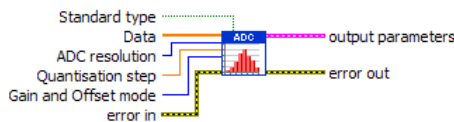
$$f_{opt} = \frac{J \cdot f_s}{M} \quad (1)$$

kde:  $J$  je celkový počet cyklov (periód) v zázname (celé číslo)  
 $f_s$  je vzorkovacia frekvencia  
 $M$  je celkový počet vzoriek v zázname

$$I \cdot J = 1 \pmod{M} \quad (2)$$

kde:  $I$  je parameter definovaný pre koherentné vzorkovanie

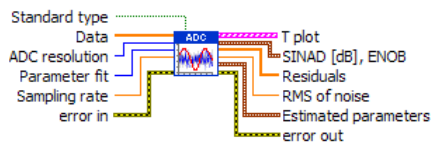
#### Histogramový test



Obr.4 Express Sinewave histogram test

Táto funkcia predstavuje samotné vyhodnotenie záznamu dát, t.j. výpočet jednotlivých kvantizačných úrovní  $T[k]$  (pre  $k=1,2,\dots,2^N-1$ ). Vytvára histogram a z neho vypočíta parametre INL a DNL podľa metódy zvoleného štandardu.

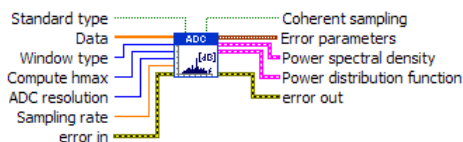
#### D. Test dynamických parametrov v časovej oblasti



Obr.5 Express Test in time domain

Táto metóda odhaľuje skreslenie výstupného priebehu AD prevodníka, ktoré je spôsobené nelineárnymi operáciami hardvéru. Aj pri tejto metóde testovania vyžadujúcej presný harmonický signál na vstupe testovaného ADP sa trochu odlišujú postupy v jednotlivých štandardoch pre získanie informácie o chybových parametroch testovaných AD prevodníkov. Okrem výpočtu troj a štvorparametrickou metódou, výpočtu chybových parametrov SINAD a ENOB, bola funkcia doplnená o výpočet SINADu pre zvolenú hodnotu vstupnej amplitúdy, podľa IEEE.

E. Test dynamických parametrov vo frekvenčnej oblasti

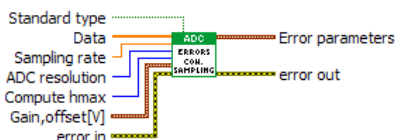


Obr.6 Express Test in frequency domain

Pri tejto metóde sa všetky chybové parametre získavajú zo spektra digitalizovaného signálu aplikovaním DFT transformácie. Pri nekoherentnom vzorkovaní sa energia harmonických zložiek rozloží aj na susedné vzorky. Preto bolo potrebné aplikovať na testovaný signál vhodnú oknovú funkciu, ktorá síce zníži hodnoty jednotlivých zložiek, ale predovšetkým po aplikovaní DFT sa rozmazanie spektrálnych zložiek na susedné vzorky zúži.

Táto funkcia pracujúca vo frekvenčnej oblasti poskytuje vstupné výpočty pre jej podriadené (volané) funkcie koherentného a nekoherentného vzorkovania opísané v ďalšom. Výber vhodného výpočtu je určený podľa vzťahu (1), ktorý na základe počtu vzoriek v zázname, vzorkovacej frekvencie a vstupnej frekvencie rozhoduje o tom, či je vzorkovanie signálu koherentné alebo nie. Pre vstup podriadených funkcií sú zo záznamu estimované parametre testovacieho harmonického signálu a to zisk  $G$ , amplitúda  $A$ , fáza  $\varphi$ , ofset  $C$  a vstupná frekvencia  $f_i$ . Pri výbere knižnice pre výpočet s nekoherentným vzorkovaním sa najprv aplikuje zvolená oknová funkcia.

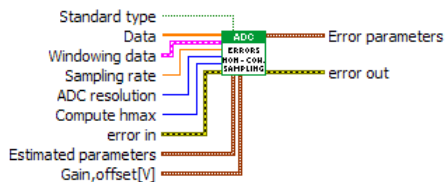
Koherentné vzorkovanie



Obr.7 Coherent sampling in frequency domain

Táto funkcia je volaná v prípade splnenia podmienok pre koherentné vzorkovanie. Samotný výstup knižnice obsahuje hodnoty vypočítaných chybových parametrov, ako NFL, SINAD, SINADfs, ENOB, SNR, THD, SFDR a amplitúdu získanú z hodnôt FFT spektra. Prahová hodnota šumu NFL, ktorá informuje o tom, kedy je ešte signál možné korektné detekovať, sa pre štandard IEEE 1241 nepočíta, lebo štandard vzorec pre výpočet NFL neuvádza. Pre správny výpočet chybových parametrov podľa projektu DYNAD, bolo nutné navyše normalizovať pred samotným výpočtom NFL, zložky FFT spektra, pričom projekt DYNAD tento fakt vo vzorci neuvádza.

Nekoherentné vzorkovanie



Obr.8 Non-coherent sampling in frequency domain

Táto funkcia je volaná v prípade nekoherentného vzorkovania. V porovnaní s koherentným vzorkovaním je doplnená o vstup oknových funkcií a estimovaných parametrov, ktoré využíva štandard IEEE 1241 k výpočtu NAD. Výstup knižnice obsahuje chybové parametre totožné s koherentným vzorkovaním. Americký štandard ani v tomto prípade neuvádza výpočet NFL. Dáta digitalizovaného signálu po aplikovaní oknovej funkcie sú rozdelené ziskom oknovej funkcie NNPG. Výpočet pre SINAD je realizovaný všeobecnou formulou, ako pomer efektívnej hodnoty amplitúdy ku efektívnej hodnote šumu a skreslenia, pretože tento štandard bližšie nešpecifikuje výpočet pre nekoherentné vzorkovanie.

Pre výpočet SNR, SINAD a NFL sú ošetrené rozmazané zložky spektra DC a základnej harmonickej. Korekciu rozšírenia spektrálnych zložiek signálu po aplikovaní oknovej funkcie



predstavuje ekvivalentná šumová šírka pásma oknovej funkcie ENBW, ktorej decibelová hodnota je pripočítaná k pomeru SNR. Celkový pomer SNR pre nekoherentné vzorkovanie je korektný až po korekcii veľkosti samotnej magnitúdy základnej harmonickej zložky. Pre túto poslednú korekciu bolo potrebné naprogramovať tzv. korekčný faktor oknovej funkcie (3), ktorý je definovaný v projekte DYNAD a jeho decibelovú hodnotu pripočítať k výpočtu pre SNR. Tento korekčný faktor napráva zníženie energie zložiek spektra po aplikovaní oknovej funkcie. Integrál, ktorým je faktor definovaný je možné aproximovať aj diskretným spôsobom [5]. Znásobením vzoriek oknovej funkcie desaťkrát a následnou aplikáciou FFT transformácie na vzorky, tak získame rozšírené spektrum oknovej funkcie. Hodnotu korekčnej vzorky v rozšírenom spektre oknovej funkcie získame súčinom chyby koherencie  $W_{10M}[\varepsilon_j \cdot x10]$ .

$$\frac{|w[0]|}{\left|W_c\left(\frac{\varepsilon_j f_s}{M}\right)\right|} \quad (3)$$

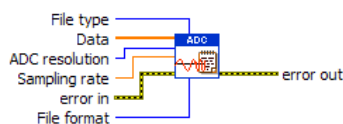
Pre výpočet SINAD je program obdobne navrhnutý ako pre SNR, ale s tým rozdielom, že menovateľ zlomku je sčítaný s korigovanými vyššími harmonickými zložkami zaokrúhlenými na celé číslo. Výber vzoriek z okna v korekčnom faktore (4) pre tieto harmonické zložky je frakciálnou funkciou zaokrúhlený na prvé desiatinné číslo (hodnoty v intervale  $[0 < r < 1]$ ). Keďže spektrum oknovej funkcie bolo rozšírené  $x10$ , nájdenú desiatinnú hodnotu  $r$  je potrebné rovnako vynásobiť desiatimi, pre správny výber korekčnej vzorky zo spektra.

$$\frac{|W[0]^2|}{\left|W_c\left(\frac{frac_r[h(J \pm \varepsilon_j)]f_s}{M}\right)\right|^2} \quad (4)$$

#### F. Zápis a čítanie dát

Posledným zložkou knižnice je funkcia umožňujúca archiváciu analyzovaných dát a ich opätovné načítanie do programu. Použité funkcie [4] boli navrhnuté podľa preddefinovaných požadovaných formátov .lvm, .xml, .tdms, .dsc.

##### Zápis dát



Obr.9 Express Write data

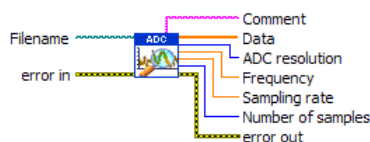
Zápisom dát detekovaných na výstupe simulovaného prevodníka, alebo načítaných zo súboru, sa vytvorí súbor vopred zvoleného typu. Vytvorený súbor obsahuje hlavičku s informatívnymi parametrami (rozlíšenie prevodníka, vzorkovacia frekvencia, počet vzoriek záznamu, frekvencia, dátum vytvorenia, meno súboru, atď.) a samotné vzorky zaznamenaného signálu.

Knižnica zápisu dát pre súborový formát .lvm, bola doplnená o dodatkové informácie a to rozlíšenie prevodníka, vzorkovaciu frekvenciu, počet vzoriek a samotné vzorky, lebo tieto informácie v základnej hlavičke chýbali. Pre zápis týchto informácií sa vygeneruje ďalší súbor spolu s pôvodne navrhnutým. Knižnica pre formát .tdms bola doplnená o rozlíšenie prevodníka a vzorkovaciu frekvenciu.

Pre formát .dsc bola doplnená knižnica o funkciu generovania dát v binárnej forme. Touto funkciou má možnosť používateľ programu archivovať dáta bez akejkoľvek straty hodnoty vzoriek. Spolu so súborom obsahujúcim hlavičku parametrov, je vygenerovaný súbor typu .dat, ktorý obsahuje samotný signál uložený v binárnej podobe. V prípade požiadavky na zápis signálu v ASCII kódach, je vygenerovaný namiesto binárneho súboru textový dokument typu .txt.

##### Čítanie dát

Funkcia pre čítanie dát volá príslušné súborové formáty detekuje formát súboru podľa prípony jeho názvu a podľa toho volá vnútorne jednotlivé rutiny pre načítanie jednotlivých formátov dát.



Obr.10 Express Read data

#### IV. DOKUMENTÁCIA K PROGRAMU

Vytvorená knižnica funkcií spĺňa požiadavky kladené na takúto knižnicu podľa štandardu LabVIEW. Preto súčasťou vytvorenej knižnice je aj dokumentácia vo forme tzv. help library, pre jednoduchšie zorientovanie sa používateľa pri používaní knižničných funkcií. Dokumentácia opisuje jednotlivé vstupy a výstupy naprogramovaných funkcií. Obsahuje aj opis implementovaného matematického aparátu včítané explicitného vyjadrenia matematických vzťahov. Je upravená podľa štýlu a systému požadovaného LabVIEW a spracovaná v anglickom jazyku.

#### V. PRÍKLAD APLIKÁCIE KNIŽNICE

Pre demonštráciu možností vytvorenej knižnice ako aj pre potreby odladenia funkcií bol vytvorený komplexný softvér zahŕňajú simulované testovanie podľa všetkých metód a postupov. Finálna verzia je umiestnená na laboratórnom web serveri <http://meas-lab.fei.tuke.sk/> a je voľne dostupná prostredníctvom internetu. Hlavná stránka obsahuje aplikáciu prezentujúcu navrhnutý softvér, ktorá umožňuje simulované testovanie AD prevodníkov. Zároveň táto stránka obsahuje referencie na príslušnú používateľskú príručku, súbor úloh pre otestovanie prevodníka vhodné pre výučbu študentov a stručnú teóriu vybraných testovacích metód.

S vopred nainštalovaným ovládačom je softvér dostupný pre širokú verejnosť na tejto internetovej stránke:

#### VI. ZÁVER

Hlavným cieľom vykonaného výskumu a vývoja bolo vytvoriť súbor softvérových funkcií – knižnicu v LabVIEW pre potreby spracovania dát z testovania AD prevodníkov. Vytvorená knižnica predstavuje komplexný nástroj, ktorý vo veľkej miere zjednoduší prácu pri vývoji programov pre testovanie jednotlivých konkrétnych AD prevodníkov a ušetrí užívateľovi veľké množstvo času spojené s programovaním zložitých matematických vzťahov ako aj s ladením a odstraňovaním chýb v takýchto programoch. Navyiac, vytvorenie demonštračného voľne prístupného programu simulujúceho testovanie AD prevodníkov umožňuje implementovať vybrané otázky z tejto problematiky do výučby ako aj propagovať výsledky práce KEMT FEI TU v Košiciach pre širšiu verejnosť prostredníctvom internetu.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (ITMS 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### POUŽITÁ LITERATÚRA

- [1] IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters, "IEEE Std 1241-2010 (Revision of IEEE Std 1241-2000)", vol., no., pp.1,139, Jan. 14 2011  
doi: 10.1109/IEEESTD.2011.
- [2] European project DYNAD – SMT4-CT98-2214, Programe Standards, Measurements, and Testing, Methods and draft standards for the Dynamic characterisation and testing of Analogue to Digital converters.
- [3] V.Vitkovic, "Pracovisko pre testovanie AD prevodníkov dostupné cez internet", Diplomová práca, Košice: TUKE-FEI 2014.
- [4] J.Priputen, "Testovanie AD prevodníkov", Diplomová práca, Košice: TUKE – FEI, 2012, 52.
- [5] Blair, Jerome J., "Selecting test frequencies for sinewave tests of ADCs," *Instrumentation and Measurement, IEEE Transactions on*, vol.54, no.1, pp.73,78, Feb. 2005  
doi: 10.1109/TIM.2004.838913
- [6] Testovanie AD prevodníkov, Šaliga, Ján, Košice : TU, 2013. 113 s., ISBN 978-80-553-1329-0

# Fuzzy riadenie SMPM

<sup>1</sup>Jaroslava ŽILKOVÁ, <sup>2</sup>Peter GIROVSKÝ, <sup>3</sup>Luboš SABOLÍK

<sup>1,2</sup> Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>jaroslava.zilkova@tuke.sk, <sup>2</sup>peter.girovsky@tuke.sk, <sup>3</sup>lubos.sabolik@tuke.sk

**Abstract** — Článok sa zaoberá návrhom a verifikáciou fuzzy vektorového riadenia synchronného motora s permanentnými magnetmi. V programe Matlab s podporou Fuzzy toolboxu je navrhnutý fuzzy regulátor rýchlosti vektorovo riadeného SMPM, ktorý je porovnaný s klasickým vektorovým riadením. Výsledky testovania týchto dvoch modelov vektorového riadenia SMPM v programe Simulink potvrdzujú, že riadenie s fuzzy regulátorom dosiahlo o trochu lepšie výsledky než riešenie s klasickým vektorovým riadením.

**Keywords** — fuzzy regulátor, synchronný motor s permanentnými magnetmi (SMPM), vektorové riadenie

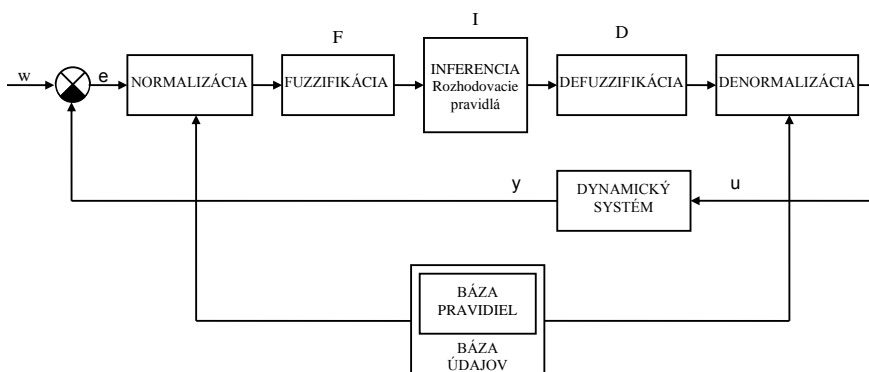
## I. ÚVOD

Článok nadväzuje na prácu [4], ktorá sa zaoberala návrhom matematického modelu synchronného motora s permanentnými magnetmi (SMPM) a jeho vektorovým riadením.

Synchronný motor patrí medzi striedavé elektrické stroje, ktorých význam rastie s rozvojom riadiacej a výkonovej polovodičovej techniky. Výhodou synchronných motorov s permanentnými magnetmi oproti budeným synchronným motorom je skutočnosť, že vďaka absencii budiaceho vinutia odpadajú budiace straty a nutnosť jednosmerného zdroja budiaceho prúdu, ako aj možnosť poruchy budiaceho vinutia.

Spôsob riadenia synchronných motorov je závislý na požadovaných vlastnostiach pohonu, pričom v súčasnosti patrí k najrozšírenejším spôsobom vektorové riadenie, ktoré bolo pôvodne navrhnuté pre asynchronný motor. Vektorové riadenie zabezpečuje prevádzku pohonu v optimálnych podmienkach v ustálenom stave i počas prechodných dejov. Princíp vektorového riadenia je založený na samostatnom riadení dvoch zložiek prúdu, pričom zmena jednotlivých zložiek je uskutočniteľná nezávisle a umožňuje riadiť zvlášť moment a magnetický tok a dosiahnuť maximálny moment v statických aj dynamických režimoch činnosti pohonu.

V posledných rokoch sa v praxi v oblasti riadenia stretávame s prístupmi, ktoré sú odklonom od klasického modelovania a ktorých spoločným znakom sú miernejšie požiadavky na presnosť opisovaných javov, ktoré nám často umožňujú použiť spôsoby riadenia blízke prirodzenému ľudskému konaniu a dosiahnuť tak v konečnom riešení často i lepšie výsledky. K týmto metódam patrí aj fuzzy logika, ktorá umožňuje v riadení systémov využiť kvalitatívne formulované skúsenosti a znalosti, ktoré máme o riešenej úlohe.



Obr. 1 Štruktúra fuzzy regulátora

Princíp fuzzy regulácie je založený na báze znalostí o systéme a rozhodovania sa pomocou pravidiel v tvare „ Ak  $e...$  – potom  $u...$ “, vytvorených zvyčajne na základe získaných skúseností. Na (obr.1) je zobrazená vnútorná štruktúra fuzzy regulátora, pričom  $w$  označuje žiadanú veličinu,  $e$  je odchýlka,  $u$  je akčná veličina a  $y$  označuje výstup systému.

Článok sa zaoberá návrhom fuzzy vektorového riadenia uhlovej rýchlosti SMPM, pričom uvádza porovnania získané modelovaním klasickej kaskádnej štruktúry riadenia SMPM s PI regulátormi a návrhom fuzzy vektorového riadenia.

## II. VEKTOROVÉ RIADENIE SMPM

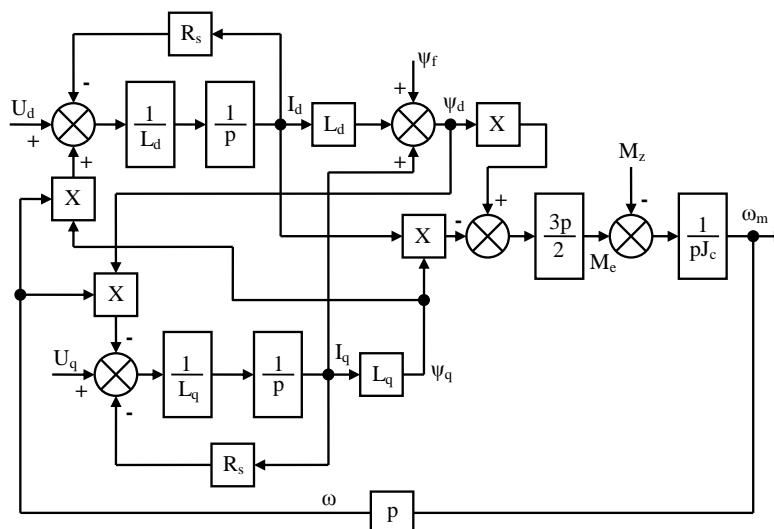
Pri vytváraní modelu SMPM sme vychádzali z nasledujúcich zjednodušujúcich predpokladov:

- neuvažuje sa vplyv nasýtenia magnetického obvodu,
- šírka vzduchovej medzery sa považuje za konštantnú, a predpokladá sa nezávislosť odporu vinutí od zmeny teploty,
- straty v železe sa zanedbávajú a mechanické straty sa pripočítavajú k záťažnému momentu,
- uvažuje sa s koncentrovanými, symetrickými vinutiami navzájom pootočenými o  $120^\circ$ ,
- zanedbáva sa vplyv vyšších harmonických,
- pre izolovaný uzol vinutia platí, že súčet prúdov je rovný nule.

Pre zjednodušenie matematického vyjadrenia sa:

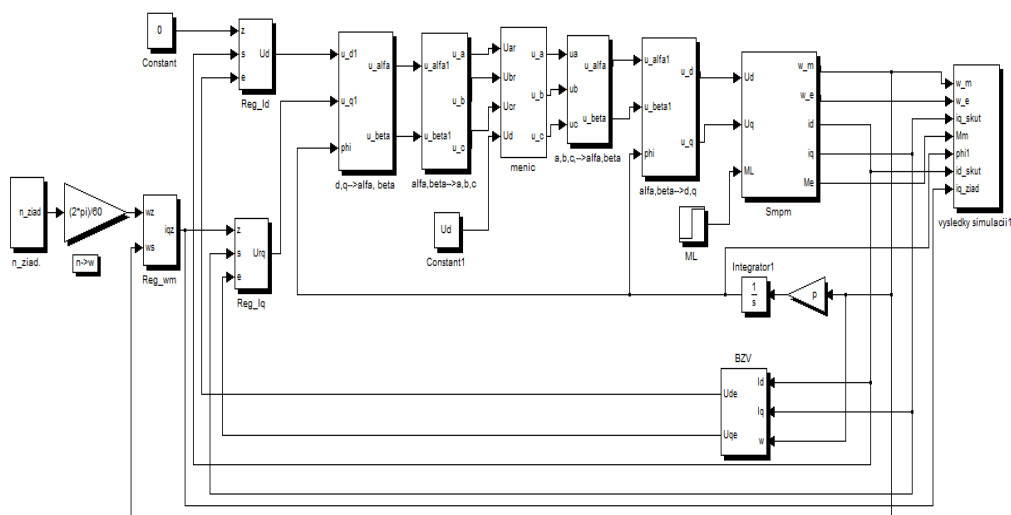
- trojfázový stroj prepočíta na dvojfázový stroj,
- viacpólový stroj prepočíta na dvojpólový, čím bude elektrická uhlová rýchlosť rovná mechanickej uhlovej rýchlosti,
- transformácia rovníc urobí do točivej súradnicovej sústavy.

Pre vyjadrenie matematického modelu striedavého synchronného stroja potrebujeme transformovať trojfázové statorové súradnice  $a, b, c$  do dvojfázových súradníc  $\alpha, \beta$  a tieto následne transformovať do rotorových súradníc  $d, q$ . Matematické rovnice platiace pre uvedené transformácie a ich spätné transformácie ako aj vyjadrenie matematického modelu SMPM sú známe [1]- [5]. Na ich základe bola v Simulinku zostavená bloková schéma matematického modelu SMPM (obr.2) a navrhnuté jeho vektorové riadenie (obr.3).



Obr. 2 Bloková schéma matematického modelu SMPM

Pri návrhu vektorového riadenia SMPM (obr.3) boli navrhnuté pre zložky statorového prúdu dva regulátory, jeden pre zložku  $i_d$  a druhý pre zložku  $i_q$  (podľa pravidla KOM) a regulátor rýchlosti (podľa KSO). Všetky regulátory boli navrhnuté podľa [5] a pre ich rozsiahlosť nie sú ich návrhy v príspevku uvedené.

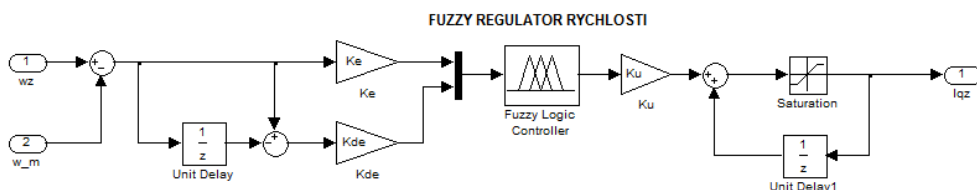


Obr. 3 Bloková schéma vektorového riadenia SMPM

Na obr.3 je zobrazená bloková schéma vektorového riadenia s reguláciou napätia a prúdu spolu s potrebnými blokmi jednotlivých transformácií súradnicových systémov [5], ktoré neboli v článku pre ich rozsiahlosť bližšie uvedené. Pred návrhom vektorového riadenia bola upravená schéma matematického modelu v súradniciach  $d, q$  z obr.2, schéma bola doplnená o zdroj napájacieho napätia a žiadanej uhlovej rýchlosti.

#### A. Návrh fuzzy regulátora

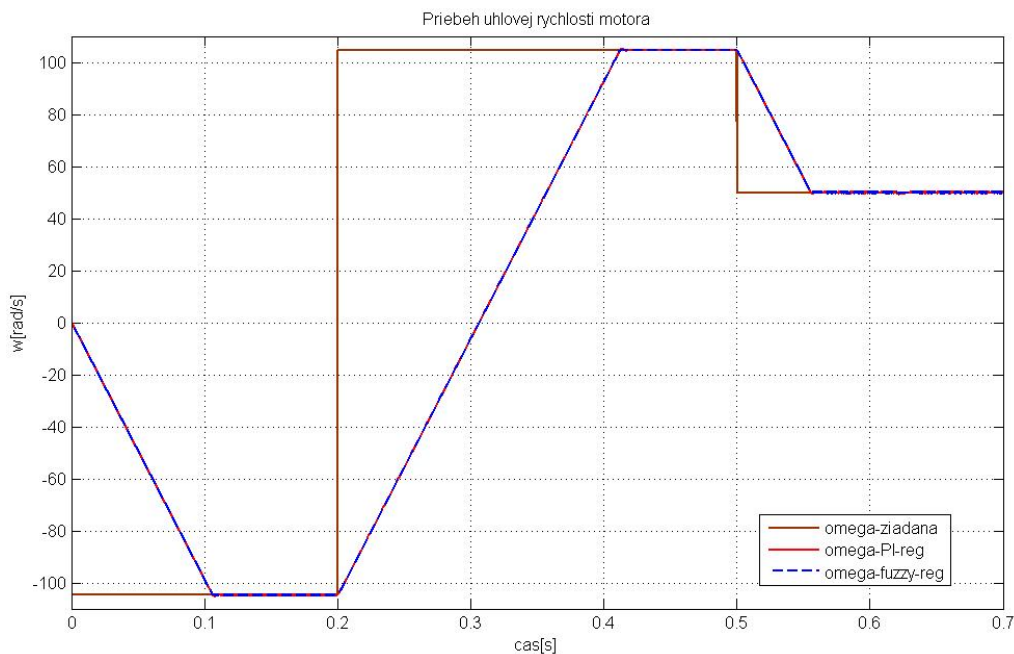
Štruktúra fuzzy riadenia SMPM bola identická s kaskádnou štruktúrou riadenia s klasickými PI regulátormi, pričom PI regulátor uhlovej rýchlosti je nahradený fuzzy PI regulátorom rýchlosti. Ladenie fuzzy regulátora rýchlosti bolo uskutočnené pomocou FIS editora, kde bolo nastavených päť hodnôt pre lingvistické premenné  $e, d_e, u$  (hodnoty lingvistických premenných  $NL, NS, Z, PS, PL$ ), pri 25 pravidlách. Pre lingvistické premenné  $e, d_e$  a  $u$  bol pre hodnoty lingvistických premenných  $NL$  a  $PL$  zvolený typ funkcie príslušnosti *trampf* a pre hodnoty lingvistických premenných  $NS, Z, PS$  bol zvolený typ funkcie *trimf*. V programe MATLAB-SIMULINK boli upravené zosilnenia  $K_e, K_{de}$  a  $K_u$  pre dosiahnutie požadovaného priebehu uhlovej rýchlosti. Štruktúra fuzzy regulátora je doplnená o obmedzenie, výstupná hodnota z fuzzy regulátora rýchlosti vstupuje ako žiadaná hodnota prúdu  $i_{qz}$  do regulátora prúdu  $i_q$ . Na nasledujúcom obrázku (obr.4) je zobrazená schéma fuzzy regulátora rýchlosti v programe MATLAB-SIMULINK.



Obr. 4 Schéma fuzzy PI regulátora rýchlosti SMPM

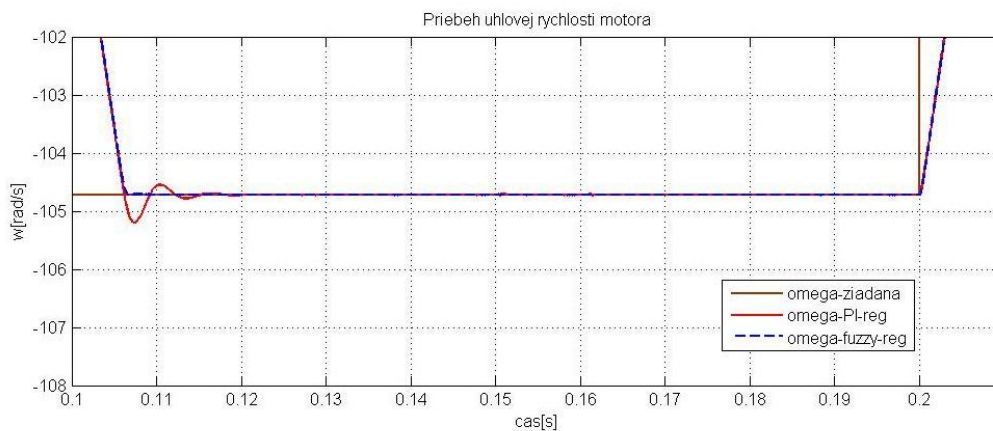
### III. SIMULÁCIE A POROVNANIE DOSIAHNUTÝCH VÝSLEDKOV

Na obr.5 je uvedený simulovaný priebeh uhlovej rýchlosti SMPM pomocou klasickej kaskádnej regulačnej štruktúry a štruktúry s fuzzy regulátorom. Žiadaná hodnota uhlovej rýchlosti bola na začiatku simulácie (obr.5) nastavená na hodnotu  $\omega_z = -104,72 \text{ rad/s}$ , v čase  $t = 0,2 \text{ s}$  na hodnotu  $\omega_z = 104,72 \text{ rad/s}$  a v čase  $t = 0,5 \text{ s}$  na hodnotu  $\omega_z = 50 \text{ rad/s}$ . Motor bol zaťažený nominálnym momentom v čase  $t = 0,62 \text{ s}$ . Pri klasickej kaskádnej regulačnej štruktúre bola ponechaná hodnota zosilnenia regulátora rýchlosti  $K_{ro}$  na vypočítanú hodnotu podľa KSO. Na obr.5 je uvedený priebeh žiadanej uhlovej rýchlosti (hnedá farba), priebeh uhlovej rýchlosti kaskádnej regulačnej štruktúry s PI regulátormi (červená farba) a priebeh uhlovej rýchlosti regulačnej štruktúry s fuzzy regulátorom (modrá farba).

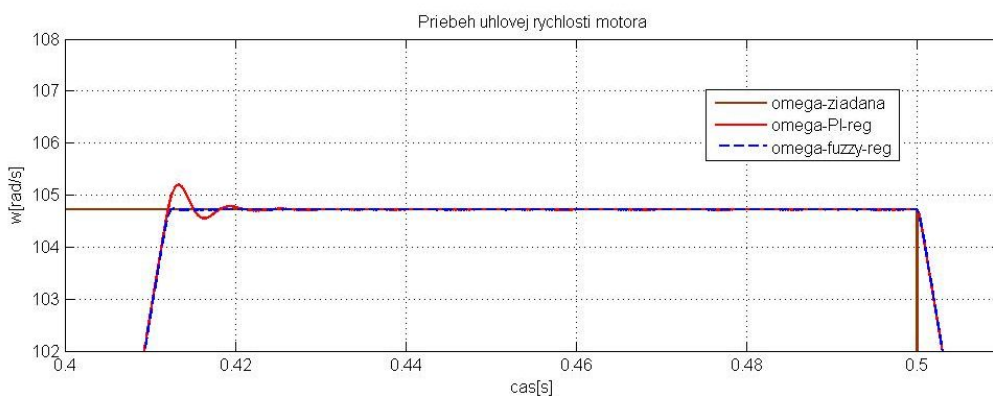


Obr. 5 Porovnanie priebehov regulácie uhlovej rýchlosti

Na obr.5 je uvedené porovnanie priebehov uhlovej rýchlosti u oboch regulačných štruktúr v celom rozsahu času simulácie. Na obr.6 je uvedený detail priebehu uhlovej rýchlosti pri žiadanej zápornej hodnote  $\omega_z = -104,72 \text{ rad/s}$  v čase  $t = 0,1 \text{ s}$  až  $t = 0,21 \text{ s}$ . Pri klasickej štruktúre s PI regulátorom dochádza k prekmitu pod žiadanú uhlovú rýchlosť na hodnotu  $\omega = -105,2 \text{ rad/s}$  a k odstráneniu regulačnej odchýlky dochádza v čase  $t = 0,123 \text{ s}$ . Pri štruktúre s fuzzy regulátorom rýchlosti nedochádza k prekmitu pod žiadanú hodnotu a rýchlosť sa ustáli na žiadanej hodnote v čase  $t = 0,11 \text{ s}$ .

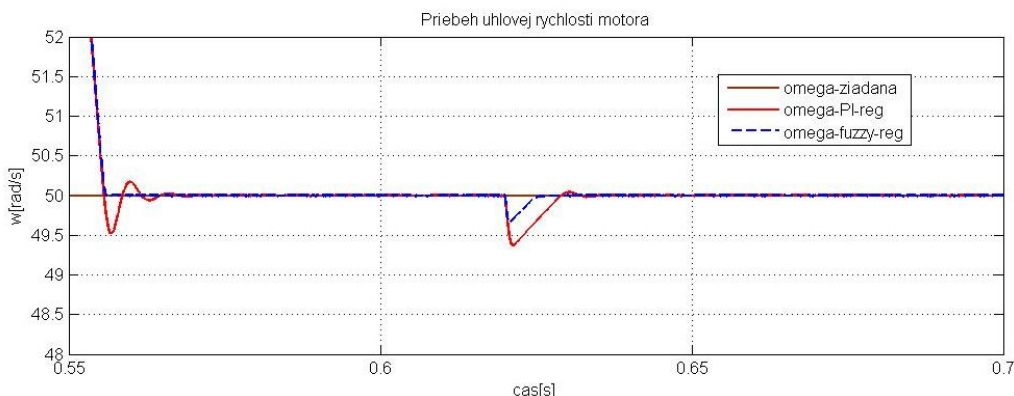


Obr. 6 Porovnanie priebehov regulácie uhlovej rýchlosti - detail 1



Obr. 71 Porovnanie priebehov regulácie uhlovej rýchlosti - detail 2

Na obr.7 je zobrazený detail priebehu dosiahnutia uhlovej rýchlosti v čase  $t=0,4s$  až  $t=0,51s$  na žiadanú uhlovú rýchlosť  $\omega_z=104,72rad/s$ .



Obr. 8 Porovnanie priebehov regulácie uhlovej rýchlosti - detail 3

Na obr. 8 je zobrazený detail priebehu uhlovej rýchlosti v čase  $t=0,55s$  až  $t=0,7s$ , kde je motor v čase  $t=0,62s$  zaťažený nominálnym záťažným momentom. Pri priebehu uhlovej rýchlosti v štruktúre s klasickým PI regulátorom rýchlosti dôjde k poklesu uhlovej rýchlosti po zaťažení na hodnotu  $\omega=49,37rad/s$  a k vyregulovaniu vzniknutej regulačnej odchýlky v čase  $t=0,64s$ . Pri priebehu uhlovej rýchlosti v štruktúre s fuzzy PI regulátorom rýchlosti dôjde k poklesu uhlovej rýchlosti po zaťažení na hodnotu  $\omega=49,7rad/s$  a k vyregulovaniu vzniknutej odchýlky v čase  $t=0,63s$ .

#### IV. ZÁVER

Dnes najrozšírenejším riadením striedavých motorov je vektorové riadenie, ktorým môžeme dosiahnuť porovnateľné, alebo dokonca aj lepšie výsledky riadenia ako u jednosmerného motora s cudzím budením, ktorý bol doposiaľ preferovaný v regulovaných pohonoch. Návrh vektorového riadenia bol prevedený pomocou kaskádnej štruktúry s klasickými regulátormi a taktiež aj s fuzzy regulátormi uhlovej rýchlosti. V článku bolo prezentované vektorové riadenie s použitím klasických regulátorov, ktoré zabezpečuje riadenie s vysokou dynamikou, pričom nevýhodou návrhu takéhoto riadenia je nutnosť poznať matematický opis riadeného systému, čo zahŕňa zložitý výpočet parametrov regulátorov. Naproti tomu výhodou návrhu štruktúry s fuzzy regulátorom je, že nie je nutné poznať matematický opis systému pre zostavenie riadenia. Pri riadení uhlovej rýchlosti SMPM dosahuje použitie fuzzy regulátorov o niečo lepšie výsledky ako použitie klasických regulátorov. Z ďalších testov (napr. pri zmene momentu zotrvačnosti motora), ktoré tu pre rozsiahlosť neboli uvedené vyplýva, že fuzzy regulátor je robustnejší oproti klasickej štruktúre, kde by bolo potrebné parametre regulátorov opäť prepočítať.

#### POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.

#### POUŽITÁ LITERATÚRA

- [1] B. K. Bose, *Modern power electronics and AC drives*. Prentice-Hall PTR, Upper Saddle River, NJ 07458 2002
- [2] I. Neborák, *Modelování a simulace elektrických regulovaných pohonů*. Ostrava, 2002.
- [3] J. Skalický, *Elektrické regulované pohony*. VUT v Brně, 2007.
- [4] J. Žilková, P. Girovský, M. Vacek, Vektorové riadenie SMPM, Electrical Engineering and Informatics IV, Proceedings of the FEI of TU of Košice, Košice, 2013.
- [5] L. Zboray, F. Ďurovský, J. Tomko, *Regulované pohony*. Viena, 2000.
- [6] P. Brandštetter, *Střídavé regulační pohony: moderní způsoby řízení*. VŠB TU, Ostrava, 1999.



# Modul pre vybrané funkcie a zariadenia fiktívnej budovy ovládané prostredníctvom Internetu

*Peter Mikluš, Ján ŠALIGA*

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky, Technická univerzita v Košiciach, Slovenská republika

peter.miklus@student.tuke.sk, jan.saliga@tuke.sk

**Abstrakt** — Článok sa zaoberá vytvorením zariadenia pre vzdialené riadenie niektorých funkcií fiktívnej budovy. Pre samotné riadenie je použitý počítač Raspberry Pi, ktorý ovláda jednotlivé moduly osadené na univerzálnej doske Humble Pi. Všetky funkcie sú spravované prostredníctvom webového rozhrania, ktoré využíva možnosti knižnice jQueryMobile. Hodnoty jednotlivých stavových veličín sú uchovávané prostredníctvom databázového servera MySQL. Ovládanie všetkých periférií pripojených prostredníctvom konektora GPIO, je zabezpečované serverovou aplikáciou napísanou v jazyku Python. Pre zabezpečenie prístupu na zariadenie v neverejných sieťach sa využíva aplikácia LogMeIn Hamachi.

**Kľúčové slová** — Raspberry Pi, Vzdialená správa, jQueryMobile, Humble Pi, LogMeIn Hamachi

## I. ÚVOD

Súčasný rozvoj komunikačnej a procesorovej techniky ako aj stále väčší dôraz na šetrenie a racionálne využitie energií zasiahol aj oblasť bývania. Toto viedlo autorou k vypracovaniu základnej štúdie s experimentálnym výstupom jednoduchého a lacného riešenia riadiacej jednotky fiktívnej budovy s minimálnymi hardvérovými nárokmi. Celý systém je realizovaný na báze známeho a obľúbeného jednodoskového počítača Raspberry PI a je zložený z dvoch hlavných častí. Prvá časť je klientske webové rozhranie, ktoré umožní užívateľovi grafické zobrazenie veličín a zmenu stavu výstupov. Druhá časť je serverová, ktorá cyklicky kontroluje všetky vstupy aj výstupy a vyhodnocuje ich. Dôležitou súčasťou je databázový server, na ktorý prístupujú obe hlavné časti a tvorí medzi nimi akýsi spojovací most.

## II. ŠPECIFIKÁCIA POČÍTAČA RASPBERRY PI A NASADENÝ SOFTVÉR

### A. Raspberry Pi Model B

Rozhodujúcim faktorom pre výber vhodného zariadenia nie je len samotný výkon, ale hlavne spotreba zariadenia a cena. Tieto kritéria dnes v porovnaní cena/výkon určujú aj ako bude zariadenie úspešné na trhu. Medzi absolútnych víťazov týchto porovnaní určite patrí vývojový modul Arduino, ktorý sa používa v tisícoch aplikácií po celom svete. Avšak modul, ktorý bol vybraný pre túto aplikáciu by som označil ako nového nástupcu Arduina. Raspberry Pi, je vývojová doska, ktorá spĺňa všetky vyššie spomenuté požiadavky. Táto doska je osadená procesorom ARM s taktom @700MHz, grafickým čipom Broadcom VideoCore IV @250 MHz, ktorý dokáže vykresľovať grafiku s rozlíšením 1080p a dekodovať video h.264/MPEG-4 AVC [1]. Je osadená dvoma USB portami štandardnej veľkosti a jedného micro USB portu, ktorý sa využíva na napájanie celého zariadenia. Pripojenie k počítačovej sieti sa realizuje pomocou štandardného LAN konektora RJ-45. Doska umožňuje priame pripojenie k monitoru prostredníctvom HDMI káblu alebo kompozitného video RCA. Zvukový výstup je vyvedený na stereo JACK konektor.

Na rozdiel od spomínaného Arduina, ktoré je osadené mikrokontrolérom s vnútornou pamäťou pre užívateľský kód. Toto zariadenie svojim výkonom, priamo umožňuje spustenie operačného systému. Tento systém sa ukladá na SD kartu, ktorá je vložená do slotu umiestneného na samotnej doske. Vo väčšine prípadov ide o modifikované distribúcie Linuxu Debian, Arch Linux ARM, Raspbian Linux, RaspBMC a ďalšie. To samozrejme dáva celému zariadeniu nový rozmer, ktorý umožňuje napr. výber programovacieho jazyka, inštaláciu



preferovaného vývojového prostredia, spúšťanie štandardných služieb ako sú webový server, ftp server, databázový server a iné. Z hľadiska komunikácie s externými zariadeniami je okrem USB portov vyvedený užívateľsky konektor pre komunikáciu s nízkoúrovňovými periférnymi zariadeniami. Tento konektor, v špecifikáciách označovaný vo väčšine prípadov ako GPIO združuje možnosť použitia rôznych rozhraní. Treba spomenúť hlavne sériové zbernice I2C, UART (RS232), SPI a takisto aj niekoľko užívateľsky programovateľných portov, vyvedených priamo z procesora.

### B. Použité softvérové technológie

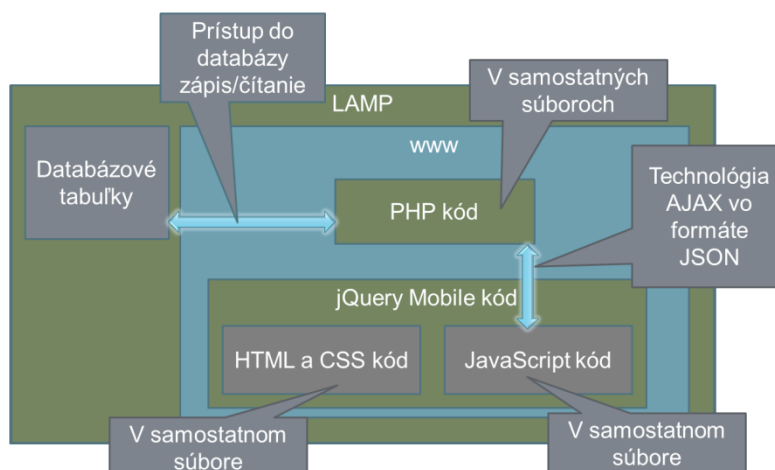
Z hľadiska využitia zariadenia ako ovládacieho centra pre rôzne spínače a snímače, je asi najdôležitejšou vlastnosťou stabilita. Preto aj operačný systém pre ostré nasadenie musí spĺňať hlavne túto požiadavku. Na výber je množstvo prevažne linuxových distribúcií prispôsobených na tento hardvér. Pre naše zariadenie sa javí ako najvhodnejší RASPBIAN Linux, ktorý vychádza z Debianu Wheezy. Čo mu zaručuje nielen stabilitu ale aj dlhodobú podporu pre opravu chýb a aktualizáciu softvéru.

Úlohu webového servera zastáva Apache HTTP Server. Na ukladanie informácií o jednotlivých stavoch je využívaný databázový server MySQL a pre bezpečné pripojenie prostredníctvom virtuálnej privátnej siete v rámci celého internetu je využívaný program LogMeIn Hamachi[5].

## III. WEBOVÉ ROZHRAINIE

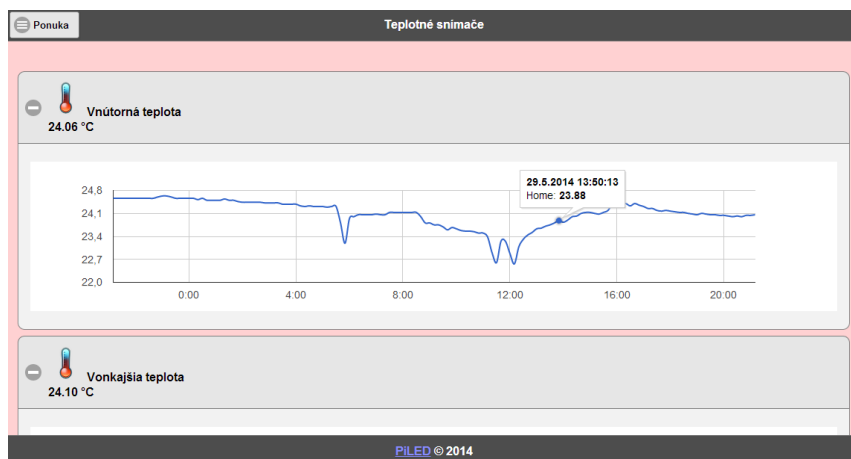
Pre vytvorenie webového rozhrania je použité jQuery Mobile. Tento softvérový rámec je určený na tvorbu užívateľsky prívetivých rozhraní, hlavne pre dnes veľmi rozšírené prenosné zariadenia. Využíva knižnicu jQuery, ktorá spája HTML a JavaScript do silného nástroja pre tvorbu interaktívneho webu. [2] Hlavne je využívaná možnosť jednoduchého vytvárania ovládacích prvkov ako sú prepínače, posuvné ovládače a rozbaľovacie zoznamy.

V základnom súbore webovej stránky index.php sú umiestnené všetky podstránky zobrazované prehliadačom. Tento prístup umožňuje prehliadaču načítať a udržiavať v pamäti kompletnú štruktúru stránky. To je dôležité hlavne pri prenosných zariadeniach, kde vďaka tomu nenastáva časové oneskorenie medzi načítavaním jednotlivých podstránok. Riadiaci kód využívajúci spoluprácu s databázou je umiestnený vo viacerých súboroch php. Prenos údajov medzi php a javascriptom je zabezpečený prostredníctvom technológie AJAX a samotný prenos je vo formáte JSON Obr. 1.



Obr. 1 - Spôsob použitia jednotlivých webových technológií

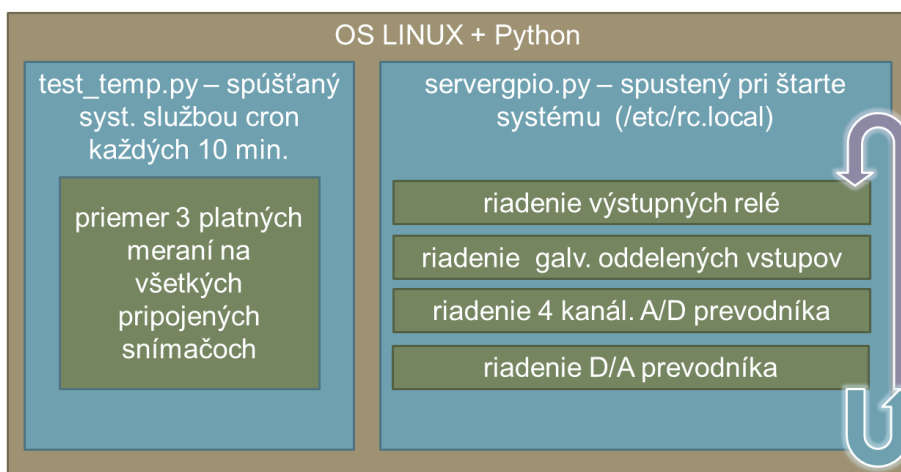
Na stránke je možné pomocou komponentu panel prepínať položky ponuky a zobrazíť tak stavy jednotlivých vstupov alebo meniť stavy výstupov. Pri každom ovládači alebo indikátore stavu je vždy zobrazený čas a dátum poslednej zmeny stavu, ktorý je pri vstupoch analógových alebo jednoduchých obnovovaný automaticky každých 5 sekúnd. Ukážka prostredia pre zobrazovanie nameraných hodnôt pomocou modulu teplotných snímačov je na Obr. 2.



Obr. 2 - Zobrazenie nameraných hodnôt

#### IV. SERVEROVÁ ČASŤ

O obsluhu všetkých modulov, ktoré sú osadené na univerzálnej doske t.j. modul reléových výstupov, modul galvanicky oddelených vstupov, modul teplotných senzorov a modul A/D a D/A prevodníkov sa stará dvojica programov napísaných v jazyku Python. Je to program *test\_temp.py*, ktorý sa stará o vyhodnocovanie teplotných senzorov a program *servergpio.py*, ktorý obhospodaruje ostatné moduly Obr. 3.



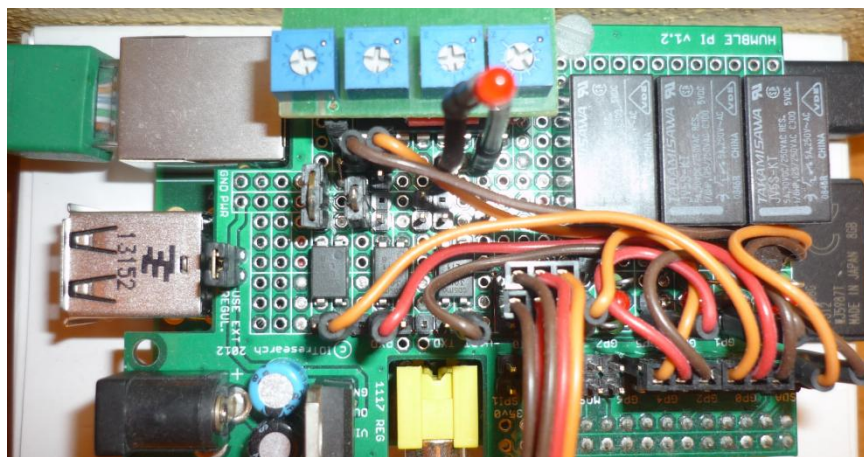
Obr. 3 - Obsluha el. modulov prostredníctvom skriptov v jazyku Python

Použitý senzory pre meranie teploty DS18B20 [6] majú základnú programovú podporu priamo v jadre Raspbian Linuxu. Je však nutné zaviesť tieto moduly do jadra, ak túto podporu chceme využívať. Viaceré tieto senzory môžu byť pripojené paralelne na jeden port GPIO4. Tento port je pri zavedení ovládacích modulov vyhradený pre tieto teplotné snímače. Moduly jadra ich na tomto porte očakávajú a nie je možné to jednoducho programovo zmeniť. Štart programu, ktorý vykonáva meranie zabezpečuje systémový démon operačného systému s názvom Cron. Tento systémový proces sa stará o periodické spúšťanie rôznych skriptov a programov.

Všetky ostatné moduly sú vyhodnocované a obsluhované programom *servergpio.py*. Konkrétne sa jedná o spínanie výstupných relé, kontrolu vstupov a riadenie A/D a D/A prevodníka prostredníctvom I2C zbernice. Tento kód vykonáva okrem ovládania hardvéru, aj čítanie a zápis do databázy.

#### V. ELEKTRONICKÉ MODULY

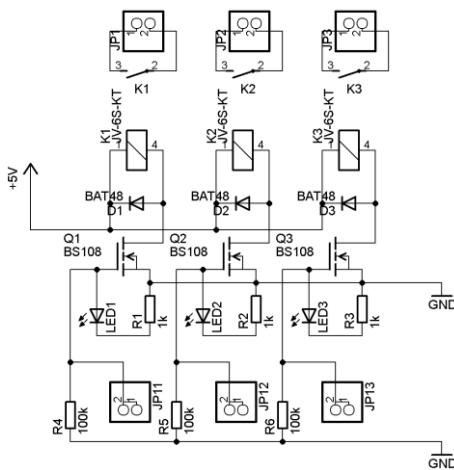
Celá hardvérová časť zariadenia bola vytvorená formou trvalých prepojení na univerzálnej prototypovej doske. Doska svojím vyhotovením kopíruje líniu samotného Raspberry Pi a vďaka tomu výsledné zariadenie pôsobí kompaktným dojmom Obr. 4. Prepojenie medzi rozširujúcou prototypovou doskou a počítačom Raspberry Pi zabezpečuje užívateľský konfigurovateľný konektor GPIO. Na tento konektor sú vyvedené nielen signály, ale aj napájacie napätia 3,3V a 5V, ktoré sú z časti využívané pre napájanie rozšírení.



Obr. 4 - RPI s pripojenou rozširujúcou doskou el. modulov

### A. Modul výstupných relé

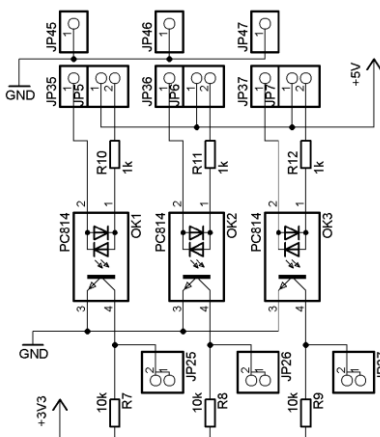
Modul umožňuje spínať prostredníctvom relé tri zariadenia. Každý výstup má v svojej vetve zapojenú signalizačnú LED diódu, ktorá oznamuje v akom stave je aktuálne konkrétny výstup. Signalizačné LED sú rozsvetované priamo signálom z procesora. Pre spínanie samotného relé je už potrebné externé napájanie, ktoré je privádzané na MOS-FET tranzistor BS108 [7].



Obr. 5 – Schéma zapojenia modulu výstupných relé

### B. Modul elektricky oddelených vstupov

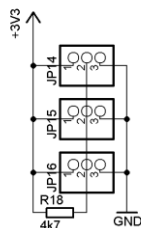
Tento modul slúži na sledovanie stavu zopnutia kontaktu alebo pripojenia napätového signálu. Má tri nezávislé vetvy a ich pracovné vstupy sú prostredníctvom optočlenu galvanicky oddelené od ostatných častí zariadenia. Keďže sa nejedná o sledovanie rýchlo sa meniacich stavov, tak sú v zapojení použité optočlenu PC814 [8].



Obr. 6 – Schéma zapojenia modulu el. oddelených vstupov

### C. Modul teplotných snímačov

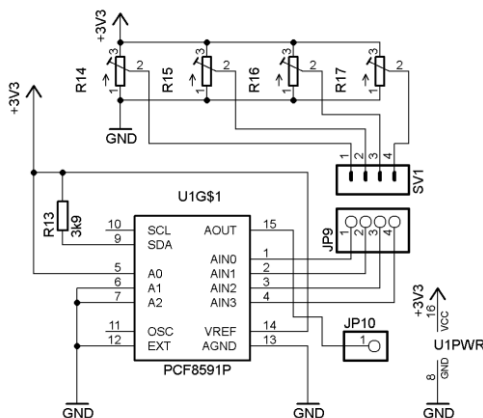
Modul v skutočnosti tvoria len konektory na pripojenie digitálnych teplotných senzorov (DS18B20) a jediný pull-up rezistor. Tieto snímače je možné pripájať paralelne, pretože každý z nich sa v systéme hlási pod svojim jedinečným kódom. Na jeden takto zapojený odpor je ich možné pripojiť až osem. Na doske sú osadené tri, ale je možné to jednoducho rozšíriť. Teplotné snímače sú napájané priamo zo systémového zdroja napätím 3,3V.



Obr. 7 – Schéma zapojenia modulu teplotných snímačov

### D. Modul A/D a D/A prevodníka

Jadrom tohto modulu je integrovaný obvod PCF8591P od spoločnosti NXP Semiconductors [9]. Na jeho čipe je integrovaný 4-kanalový A/D prevodník s rozlišovacou schopnosťou 8-bitov a jednocanálový D/A prevodník schopný meniť výstupné napätia v rozsahu taktiež 8-bitov. Maximálne výstupné napätie D/A prevodníka je určené napätím pripojením na vývod VREF a AGND. V našom prípade je pripojený priamo na napájacie napätie 3,3V. V dôsledku nízkeho odberu tohto obvodu je pripojený priamo na zdroj Raspberry Pi a preto môže pracovať aj bez externého napájania. Taktiež však tento obvod môže pracovať aj z 5V signálovými úrovňami pri napájaní 5V. Obvod komunikuje prostredníctvom dvojvodičovej sériovej linky I2C. Na vstupy jednotlivých kanálov A/D prevodníkov je pomocou konektora pripojený plošný spoj so štyrmi odporovými trimrami. Slúžia len na odskúšanie funkcie prevodníkov. Na výstup D/A prevodníka je taktiež na predvedenie a odskúšanie funkcionality pripojená pomocou konektorov LED dióda, ktorá zmenou svojho jas signalizuje zmenu napätia výstupu.



Obr. 8 - Schéma zapojenia modulu A/D a D/A prevodníka

## VI. ZÁVER

Systém navrhnutý v štúdiu bol realizovaný a odskúšaný v prevádzke vo virtuálnej budove. Napriek svojej jednoduchosti si vhodným pripojením snímačov a aktuátorov môže toto zariadenie nájsť svoje uplatnenie v objektoch, ktoré sú využívané skôr sporadicky ako napr. záhradné chaty, chalupy a iné rekreačné budovy. Pre takéto objekty by ho bolo vhodné doplniť o GSM modul a záložný zdroj, ktoré by mu umožnili plniť aj zabezpečovaciu úlohu.

Všetky softvérové riešenia použité v tejto práci, sú voľne dostupné a okrem programu na vytvorenie virtuálnej privátnej siete majú aj otvorený zdrojový kód. Toto sa vlastne odvíja od základnej myšlienky projektu Raspberry Pi o priblížení programovania širšej verejnosti najmä mládeži.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (ITMS 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### POUŽITÁ LITERATÚRA

- [1] The Raspberry Pi Foundation, „<http://www.raspberrypi.org/>“, 2008. [Online]. Available: <http://www.raspberrypi.org/>. [Cit. 04 03 2014].
- [2] J. Vávrů, jQuery Mobile, Brno: Computer Press, 2013.
- [3] M. Ponkrác, PHP a MySQL bez předchozích znalostí, Brno: Computer Press, a.s., 2007.
- [4] D. Harms a K. McDonald, Začínáme programovat v jazyce Python, Brno: Computer Press, a.s., 2008.
- [5] hamachi.cz, „Oficiální český web hamachi.cz“, [Online]. Available: <http://www.hamachi.cz/>. [Cit. 05 2014].
- [6] Maxim Integrated Products, Inc., „<http://www.maximintegrated.com/>“, 2008. [Online]. Available: <http://datasheets.maximintegrated.com/en/ds/DS18B20.pdf>. [Cit. 04 03 2014].
- [7] Philips Electronics N.V., „Datasheet BS108 N-channel D-MOS transistor“, 08 05 2001. [Online]. Available: [http://www.nxp.com/documents/data\\_sheet/BS108.pdf](http://www.nxp.com/documents/data_sheet/BS108.pdf). [Cit. 01 03 2014].
- [8] SHARP CORPORATION, „Datasheet PC814 PHOTOCOUPLER“, 23 01 1998. [Online]. Available: <http://www.alldatasheet.com/datasheet-pdf/pdf/43367/SHARP/PC814.html>. [Cit. 01 03 2014].
- [9] NXP B.V., „Datasheet PCF8591 8-bit A/D and D/A converter“, 23 06 2013. [Online]. Available: [http://www.nxp.com/documents/data\\_sheet/PCF8591.pdf](http://www.nxp.com/documents/data_sheet/PCF8591.pdf). [Cit. 02 2014].

# Photorealistic images execution in parallel environment

<sup>1</sup>*Branislav SOBOTA*, <sup>2</sup>*Martin KEREKÁČ*, <sup>3</sup>*Štefan KOREČKO*

<sup>1,2,3</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>branislav.sobota@tuke.sk, <sup>2</sup>martin.kerekac@student.tuke.sk, <sup>3</sup>stefan.korecko@tuke.sk

**Abstract** — Ray tracing is a compute-intensive method of 3D scenes rendering with photorealistic outcomes. To increase its speed of calculation distributed computing is used. To balance the load methods for scene evaluation and partitioning are used. A system for distributed calculation of ray tracing is created. To control the workflow the Hadoop platform is used, a cluster platform and framework for creation of MapReduce tasks. The cluster based on this system can scale up to thousands of nodes. For rendering POV-Ray, a popular ray tracer, is used.

**Keywords** — ray tracing, parallel execution, distributed execution, Hadoop

## I. INTRODUCTION

Ray tracing is a method of rendering 3D scenes based on a mathematical model of light reflectance.[1] In this method the interaction of light and objects in the scene is traced. Unlike the real world, this method does not trace light rays emitted from light sources, but the light rays outgoing from the camera. For each pixel of the resulting image a ray is sent. By doing this interactions with the rays that would never come into the camera are not computed. This helps to reduce computational complexity of the process.

By this point ray tracing method is consistent with the method of ray casting. The difference occurs when the beam collides with the first object. In ray casting the color value for the pixel is calculated and it is continued with the calculation. In the ray tracing method when a collision with an object is detected, the phenomenon of reflection, refraction, or other optical phenomena included in the calculation model, occurs. Thus new beams are recursively created and traced. This can theoretically continue indefinitely, in practice, however it is limited, either by recursion depth or total distance traveled by the rays. The final color of pixel is computed up from all the rays. This process is computationally intensive, but it offers the possibility of parallelization, because the calculation of each color pixels is independent from the calculations of other pixels. [2]

Rendering using ray tracing offers several programs. One of widely used programs is POV-Ray. [3]. It offers a wide range of options and has open source code, which makes it a suitable candidate for the research activity and the results would be applicable in practice.

As regards the means of supporting parallel computations several options are offered. The first is an implementation of the MPI interface. [4] The advantage of this interface is a high degree of calculation control. We can control what is calculated on what computer and also the implementation in C language ensures high speed calculations. Perhaps the biggest drawback is that high control of distribution of calculations is only possible because of the need to manually write the control code. Some solutions for parallel ray tracing based in MPI already exist. They can be found in [5] and [6].

Another alternative is the Hadoop framework. [7] It is a system originally designed to map-reduce data processing. Its advantages are its own distributed file system that takes care of sending and collecting data. The system also takes care of distributing the actual handler code written by programmers. So when editing the program, are changes automatically reflected on all cluster nodes.

Another problem is to balance load evenly between all nodes of the cluster. It is necessary to

evaluate the scene and identify the complexity of its parts. Some methods are presented in [8]. In this work we use a method based on counting of intersections of rays, casted for each pixel, with objects.

## II. SCENE EVALUATION

The scene evaluation algorithm is based on the actual ray tracing algorithm. The overview of the ray tracing algorithm is shown on next pseudo-code fragment:

```
For each pixel in the image
{
  Create a ray pointing from the eye through the pixel in the
  rendering plane
  Initialize nearestT to infinity and the nearestObject to None

  For each object in the scene
  {
    If a beam intersects the object
    {
      If the distance of the intersection is less than
      nearestT
      {
        Set nearestT to intersection distance
        Set the nearestObject to this object
      }
    }
  }

  If the nearestObject is equal to None
  {
    Fill the pixel with background color
  }
  Otherwise
  {
    Calculate the pixel color
  }
}
```

The scene is evaluated by number of intersections that need to be computed for each ray with all objects in the scene. From the ray tracing algorithm, it may seem that this number is the same for all pixels. After all, the intersection is calculated with each object, for each ray. But it's not quite true. Although it is calculated, how many objects the ray intersects and then the nearest intersection is chosen, but the first count intersections with the so-called bounding objects. These geometric primitives, usually cubes, spheres and cylinders, which are the object described (in the geometric sense). Only when it is found that the ray intersects the boundary object, the intersection with the actual object is calculated. It is easier to calculate intersection of a ray with a cube, where is intersection line with several planes, than to calculate intersection of a ray with a torus which is given by a quartic equation. This will save a large amount of calculations for objects not intersected by a ray. The actual algorithm is shown on flow chart in Fig. 1.

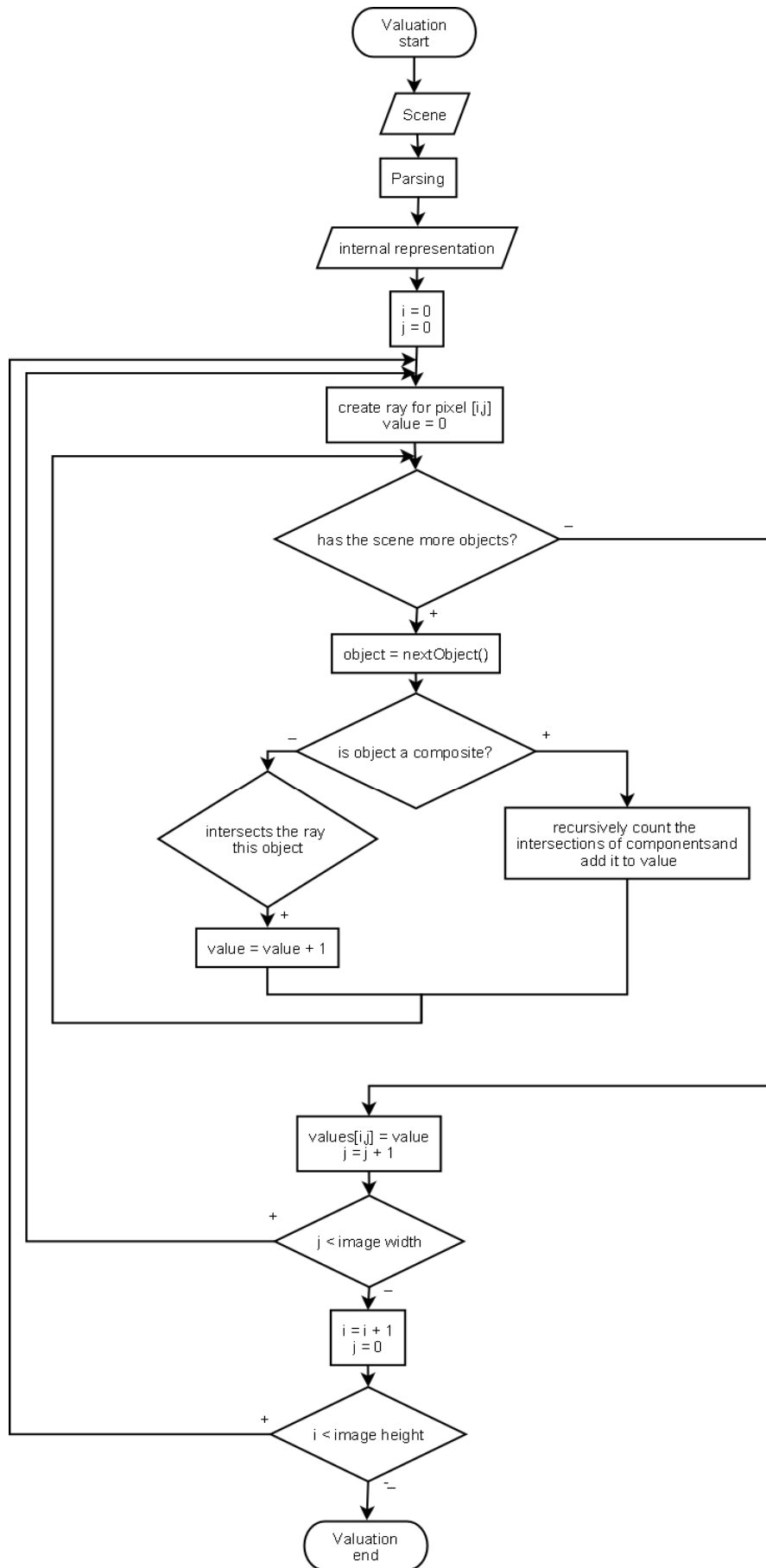


Fig. 1 Scene evaluation algorithm flow chart



### III. SCENE PARTITIONING

When the scene is evaluated, it must be divided into equally valued sections. Also there exist a number of methods. Some were presented in [6].

In this work, the sections are created by adding columns of pixels one pixel wide. In this procedure, it is necessary to first calculate the total valuation of the whole picture. This is divided by the number of sections. This number indicates the target valuation of each section. It is now possible to proceed to the very creation of the section. The process starts from the left end of the previous section, the first since from the first column. This will be the basis of the first section. To the first column are gradually added more columns. After adding each column the overall valuation of the section is calculated. If the overall valuation is equal to or greater than the target valuation of a section, is started with the creation of the next section. The algorithm is shown on the next pseudo-code fragment and flow chart in Fig. 2.

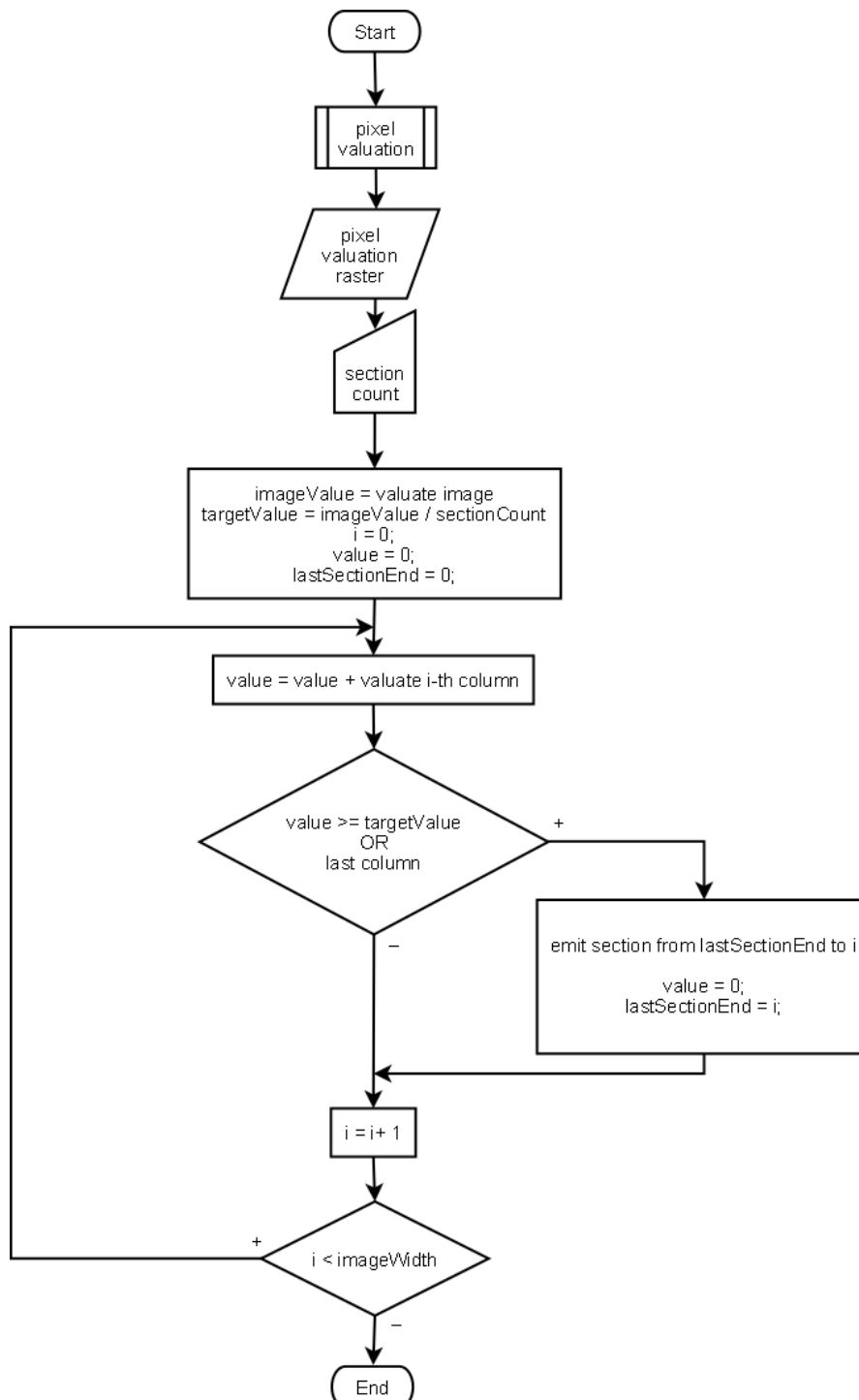


Fig. 2 Scene partitioning algorithm flow chart

```
partitions = empty list
targetValuation = imageValuation / section count
lastEnd = 0
valuation = 0
i = 0

while i < imageWidth do
{
    valuation = valuation + (valuation of the i-th column)

    if valuation >= targetValuation OR i = imageWidth - 1 do
    {
        Add section from lastEnd to i to sections
        valuation = 0
        lastEnd = i
    }

    i = i + 1
}
```

#### IV. CLUSTER NODE VALUATION

Some solutions [9] are available on internet. This work presents yet another solution focused on Hadoop clusters. It is created without dependencies other than the Hadoop framework.

The method used for valuation of nodes is based on arithmetic operations for benchmark the CPU. Array operations are incorporated to benchmark the memory. The arithmetic and array operation are repeated several times. The value evaluating the node is the execution time of this cycle. The smaller this value is the better.

#### V. CONCLUSION

A functioning application was produced, enabling distributed computing of ray tracing on the Hadoop platform. It enables statical load balancing based on the complexity on computation difficulty of the scene sections. It evaluates the scene en partitions the scene based on this evaluation.

Part of produced application can also evaluate the cluster nodes of a Hadoop cluster.

#### ACKNOWLEDGMENT

This work is supported by KEGA grant project No. 050TUKE-4/2012 “Application of Virtual Reality Technologies in Teaching Formal Methods” and this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF.(20%)

#### REFERENCES

- [1] K. Suffern. Ray tracing from the ground up. Mirage, 2007.
- [2] A. Chalmers, T. Davis, and E. Reinhard. Practical parallel rendering. AK Peters, Ltd., 2002.
- [3] <http://www.povray.org/>, august 2010
- [4] Foster, Ian. Designing and Building Parallel Programs. Boston: Addison-Wesley, 1995.
- [5] Okoš, Róbert: Main computing part of distributed system for realistic rendering by Raytracing method in network environment. Košice: Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic, 2009.
- [6] Rusyniak, Peter: Photorealistic method for large graphical data-set processing. Košice: Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic, 2008.
- [7] <http://hadoop.apache.org/>, august 2010
- [8] Danková, Eva: Preparation part of distributed system for fotorealistic rendering by Raytracing method in network environment. Košice: Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic, 2009.
- [9] <http://mkweb.bcgs.c.ca/clusterpunch/>, august 2010

# Lygon - a programming language for linear logic

<sup>1</sup>Emília DEMETEROVÁ, <sup>2</sup>Zuzana DORKINOVÁ, <sup>3</sup>Daniel MIHÁLYI,  
<sup>4</sup>Valerie NOVITZKÁ

<sup>1,2,3,4</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>emilia.demeterova@tuke.sk, <sup>2</sup>zuzana.dorkinova@student.tuke.sk,  
<sup>3</sup>daniel.mihalyi@tuke.sk, <sup>4</sup>valerie.novitzka@tuke.sk

**Abstract** – Linear logic is a logic with an expressive power what other logics do not own because it contains two type of conjunction and disjunction and other connectives which other logics do not. Lygon is a logical programming language and since its implementation is not very used. Therefore it is very difficult to understand it. The aim of this contribution is to show that exists a programming language which is suitable for linear logic. By this programming language is possible to define linear parameters, so applicable only once.

**Keywords** – Linear logic, logical programming, Lygon

## I. INTRODUCTION

Linear logic belongs among newer logic. Linear logic has a very great expressive power. With linear connectives which were defined by Girard [3], [4], linear logic is able to describe complex processes not only from the real life but from computer science too. It can express internal a external non-determinism, parallelism, unlimited resources, dynamics of processes, the construction of *xor* in programming language. By linear implication  $\multimap$  we can express the dynamic of processes, sequentiality and causality of processes [8]. Linear implication disposes with a power which can determine which process succeeds after previous. Linear formulae expressed by linear logic expresses resources or actions [7].

Another important feature of linear logic is its ability handle with resources. Resources, which linear logic uses are the most important resources, i.e. logical time and logical place in programming too.

For a real implementation of linear logic exists a linear logic programming language in which is ability to form examples from linear logic [2], [6].

## II. LYGON

Lygon is a linear logical programming language which extends logical programming language Prolog with a features of linear logic [1].

Lygon is a work of J. Harland and D. Pymbut who designed the language, M. Winikoff implemented it and Yi Xiao Xu is an author of a debugger [5].

### A. Theoretical foundations of the logic programming

Literal is an any atomic formula  $p$  which is positive  $p$  or negative  $\neg p$ . Clause is any disjunction of literals. Horn clause is a clausula containing at most one positive literal (the others are negative). Horn formula is a formula which is a conjunction of Horn clauses, e.g.

$$\neg p \vee \neg q \vee \dots \vee \neg t \vee u$$

Horn clauses are used to improve the computation of answers to a specific user's questions and are written up as an implication  $(p \vee q \vee \dots \vee t) \Rightarrow u$ .

We show it on an example. We rewrite the Lygon's rule

`checkMul(A,B) <- lt(A,0) & lt(B,0) & is(C,0-A) * is(D,0-B) * multiplying(C,0,D).`

from the example **II-D1** to a Horn clause. This rule has the form as

$$lt(a,0) \& lt(b,0) \& is(c,0-a) \otimes is(d,0-b) \otimes multiplying(c,0,d) \multimap checkMul(a,b)$$

in linear logic.

We used DeMorgan rules to rewriting the linear logic's form on Horn clauses.

$$lt(a, 0)^\perp \otimes lt(b, 0)^\perp \otimes is(c, 0 - a)^\perp \wp is(d, 0 - b)^\perp \wp multiplying(c, 0, d) \wp checkMul(a, b)$$

In this form is shown that the Horn clause contains only one positive literal  $checkMul(a, b)$ .

### B. The syntax of Lygon

The syntax of Lygon is early the same with the syntax of Prolog [9]. The difference between them are that the body of clauses are not presented as a sequence of atoms but as a formulae of linear logic [11], [12]. The program clauses are reuseable, i.e. non-linear. But the user may define it as linear by putting the *linear* prefix before the clause.

The syntax of Lygon is given by BNF grammar.

$$\begin{aligned} G &::= G * G | G @ G | G \& G | G \# G | !G | V exist G | once G | ?neg D | one | bot | top | A | neg A \\ D &::= [linear] A_1 \# \dots \# A_n < -G. | [linear] A. \\ A &::= A \text{ Prolog atom} \\ V &::= A \text{ lower-case symbol (eg. "x")} \end{aligned} \tag{1}$$

where [x] means that "x" is optional, non-terminals are written by *italics* and terminals by teletype.

The connectives from linear logic are translated into Lygon as:

$$\begin{array}{lll} \otimes \Rightarrow * & \& \Rightarrow \& & \wp \Rightarrow \# \\ \oplus \Rightarrow @ & \multimap \Rightarrow - > & \perp \Rightarrow neg- \\ 1 \Rightarrow one & \top \Rightarrow top & \perp \Rightarrow bot \end{array}$$

### C. Built-in predicates

Lygon involves some predicates which are useful for metaprogramming, arithmetical and input-output operations [10]. They are in the form: the predicate's name/the number of arguments.

- **print/1**: predicate is used to print arguments;
- **nl/0**: predicate will go to a new line;
- **input/1**: allows to read Prolog's terms and bind them to the argument;
- **system/1**: through this predicate argument shall be transmitted to the system where has been executed as a command;
- **is/2**: this predicate bounds the result on its first argument and the second argument rates as an arithmetic formula, e.g.  $is(X, 5*6) * print(X)$ . puts to  $X$  the value  $5*6$  and print it out;
- **lt/2**: compares the arguments it contains. Arguments must be numerical and the first must be less than the second;
- **builtin/1**: includes a built-in predicate, which must be equal to its argument;
- **readprog/1** and **readgoal/1**: these predicates allow reading terms and their conversion to Lygon's programs and goals, e.g.  $readgoal /1$  on  $a\#b*c$  returns  $atom(a)\#atom(b)*atom(c)$ ;
- **prolog/1**: to perform gives an argument to basic Prolog's system if the programmer decides to extend the basic set of builtin predicates;
- **call/1**:  $call(X)$  is the same as  $X$ .

### D. Some examples in Lygon

The next examples show how are programs written in Lygon.

On the first examples we will show how are integers multiplying in Lygon.

#### 1) Multiplying integers:

```
mul <- print('Function to multiplying') * nl * print('Please, enter
the numbers to multiply, e.g. mul(4,3)') * nl.
mul(A,0) <- print('Expression ') * print(A) * print('x0 is equal to
0') * nl.
mul(0,A) <- print('Expression 0x') * print(A) * print(' is equal to
0') * nl.
mul(A,B) <- print('Expression ') * print(A) * print('x') * print(B) *
print(' is equal to ') * checkMul(A,B).
checkMul(A,B) <- lt(A,0) & lt(B,0) & is(C,0-A) * is(D,0-B) *
multiplying(C,0,D).
checkMul(A,B) <- lt(B,0) & multiplying(B,0,A).
checkMul(A,B) <- multiplying(A,0,B).
```

```
multiplying(A,B,0) <- print(B) * nl.  
multiplying(A,C,B) <- is(D,B-1) * is(E,A+C) * multiplying(A,E,D).
```

The user enter the `mul(A,B)` expression to multiplying integers  $A$  and  $B$ .

First the program checks that one of the number is equal to zero. If yes than the program prints zero. If no, is called a `checkMul` function which allows to multiply negative numbers too. At the begining compare both numbers with a zero and if they are negative creates their absolute value which is in the step:

```
checkMul(A,B) <- lt(A,0) & lt(B,0) & is(C,0-A) * is(D,0-B) *  
multiplying(C,0,D).
```

Multiplying numbers take place in the function `multiplying(A,C,B)` in which the parameters  $A,B$  are input numbers,  $C$  is a parameter for an intermediate data. In the first step the number  $B$  is decremented and its value is memorized into the auxiliary variable  $D$ . Into an auxiliary variable  $E$  is inserted the sum of variable  $A$  and a continuous multiplication result and recursively called the same function with new parameters- adding `(A,E,D)`. This step is repeated until the variable containing the second input integer is decremented at zero. Than the function `adding(A,B,0)` is called and gives the result.

The program multiplies the integers up to a maximum value of the product of 536 870 911.

### 2) *The square of the integer:*

```
pow <- print('The square of the integer') * nl * print('Please,  
enter the number to square, e.g. pow(4)') * nl.  
pow(0) <- print('The square of the integer 0 is 0') * nl.  
pow(A) <- lt(A,0) & print('The square of the integer ') * print(A)  
* print(' is ') * is(B,0-A) * multiplying(B,0,B).  
pow(A) <- print('The square of the integer ') * print(A) *  
print(' is ') * multiplying(A,0,A).
```

The user enter the `pow(A)` expression to squaring integer  $A$ .

First the program checks if the integers the users wants to squaring is equal to zero. If yes, printing the result. If no, the program checks that the integer is positive or negative. If the integer is negative, it creates its absolute value and the function `multiplying` is called. In the case that the integer is positive, the function `multiplying` is called directly and the result is printed.

The program squaring the integer up to a maximum value of the power exponent of 9374. It is possible to squaring higher integer too, but the program does not work correctly. It is also necessary that the `multiplying` function is a part of the source code of the program.

### 3) *Fibonacci number:*

```
fibonacci <- print('Fibonacci number') * nl * print('Please,  
enter how many item of the sequence is to be printed (max 43),  
e.g. fibonacci(5)') * nl.  
fibonacci(A) <- print('Fibonacci number up to') * print(A) * print  
('. item') * nl * num(1,1,A,0).  
num(1,1,A,B) <- lt(B,A) & print(1) * nl * is(C,B+1) * num(1,2,A,C).  
num(X,Y,C,D) <- (lt(D,C) & lt(X,Y)) & is(A,X+Y) * is(B,Y+A) * print  
(X) * nl * is(E,D+1) * lt(E,C) & (print(Y) * nl * is(F,E+1) *  
num(A,B,C,F)).
```

The user after the entering the `fibonacci(A)` expression receives the Fibonacci sequence of numbers up to  $A$ . item.

In the Fibonacci sequence of numbers every next one member is the sum of the two previous. The members of the sequence are called Fibonacci numbers- 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233 etc.

The program calls as first the `num(A,B,C,D)` function, where  $A$  is a current number of the sequence,  $B$  is a next number of the sequence,  $C$  is the last row of the Fibonacci sequence,  $D$  is a current value of the row.

Because the sequence is starting with the numbers 1,1 it is created a `num(1,1,A,B)` function which first compares the variables  $A$  and  $B$ . If  $A$  is larger number, the program writes '1', the  $B$  value increments and is inserted into the  $C$  variable and then the `num(1,2,A,C)` function is called.

The principle of creating the sequence is realized in `num(X,Y,C,D)` function. The current sequence number must be less than the following and simultaneously the current row must be less than the maximal sequence row. If this conditions are fulfilled the other two next sequence

numbers are created and the actual number is printed. The actual sequence row increments and if this row is still less than the maximal, the next number of the sequence is printed.

Recursion of this function is performed until the actual sequence row will not be the same as the maximal so the row given by user.

Program works correctly until the 43.sequence row. The last correctly displayed number is 433 494 437.

### III. CONCLUSION

This contribution gives a brief overview of linear logic and a linear logic programming language-Lygon. It shows how is linear formula from linear logic rewritten on the clause of Lygon. It contains also some examples written in it. Examples are algebraic, but in the future research we will concern on designing examples in Lygon from linear logic mostly oriented on component-oriented programming.

### REFERENCES

- [1] J.-marc. Andreoli, "Logic Programming with Focusing Proofs in Linear Logic," *Journal of Logic and Computation*, vol. 2, pp. 297–347, 1992.
- [2] E. Demeterová, D. Mihályi and V. Novitzká, "Logical programming in informatics," *In:Electrical Engineering and Informatics 4 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Koice. - Koice : TU.*, pp. 867–872, ISBN 978-80-553-1440-2, 2013.
- [3] J.-Y. Girard, "Linear Logic,"*Theoretical Computer Science*, vol. 50, pp. 1–102, 1987.
- [4] J.-Y. Girard, "Linear Logic: Its syntax and semantics,"*Cambridge University Press*, 2003.
- [5] J. Harland, D. Pym and M. Winikoff, "Programming in Lygon: an overview," *Proceedings of the Fifth International Conference on Algebraic Methodology and Software Technology*, pp. 391–405, 1996.
- [6] V. Novitzká and D. Mihályi, "Resource-oriented Programming Based on Linear Logic," *Acta Polytechnica Hungarica*, 4(2);pp. 157–166, 2007.
- [7] V. Novitzká, D. Mihályi, V. Slodičák, "Linear logical reasoning on programming," *Acta Electrotechnica et Informatica No.3*, vol. 6, 2006.
- [8] D. Mihályi and V. Novitzká, "What about Linear Logic in Computer Science?," *Department of Computers and Informatics, Technical University of Koice, Acta Polytechnica Hungarica*, vol. 10, No. 4, 2013.
- [9] J. Harland, D. Pym and M. Winikoff, "Programming in Lygon: A Brief Overview," *Logic Programming:The 1995 International Symposium*, pp. 636, MIT Press, ISBN:9780262291248, 1995.
- [10] M. Winikoff, "Hitch Hikers Guide to Lygon 0.7," *Department of Computer Science, The University of Melbourne Parkville, Victoria 3052 Australia*, 38pp, 1996.
- [11] M. Winikoff and J. Harland, "Implementing the Linear Logic Programming Language Lygon," *In Proceedings of the 1995 International Logic Programming Symposium, Portland, Oregon*, pp. 66–80, 1995.
- [12] M. Winikoff and J. Harland, "Some Applications of the Linear Logic Programming Language Lygon," *Proceedings of the 19th Australasian Computer Science Conference, Melbourne, Australia*, 1996.

# Analýza trhu s elektrinou

<sup>1</sup> Alexander MÉSZÁROS, <sup>2</sup> Marek PAVLÍK, <sup>3</sup> Roman TÓTH

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup> alexander.meszaros@tuke.sk, <sup>2</sup> marek.pavlik@tuke.sk, <sup>3</sup> roman.toth@student.tuke.sk

**Abstrakt** — Liberalizovaním energetického odvetvia sa vytvorilo konkurenčné prostredie, ktoré prináša nové možnosti, ale zároveň aj nové riziká podnikania. Článok analyzuje možnosti pri kúpe a predaji elektrickej energie na česko-slovensko-maďarskom trhu s elektrinou.

**Kľúčové slová** — liberalizácia, cena elektriny, účastníci trhu s elektrinou, obchodovanie s elektrinou

## I. ÚVOD

Elektrina je základom národného hospodárstva každého štátu a to vďaka jej nenahraditeľnosti v koncovom použití (informačné a telekomunikačné zariadenia aj rôzne využitie v priemysle). Jej spotreba stále rastie. Odvetvie zaoberajúce sa výrobou, prenosom, distribúciou a spotrebou elektrickej energie sa nazýva elektroenergetika. Elektroenergetický systém predstavuje technologický proces zahrňujúci výrobné zdroje elektrickej energie, prenosovú a distribučnú sústavu a zariadenia pre konečnú spotrebu elektrickej energie. Jeho cieľom je dodávka požadovaného množstva elektrickej energie smerom k odberateľom v danej dobe v dohodnutom množstve a kvalite, s minimálnym vplyvom na životné prostredie. Optimalizácia investičných a prevádzkových nákladov je podmienkou úspešnosti v obchodnom prostredí konkurenčných spoločností [1].

## II. LIBERALIZÁCIA A ÚČASTNÍCI TRHU S ELEKTRINOU

V celej Európe pred začiatkom liberalizácie na začiatku 90-tých rokov boli dodávky sprostredkované v danej geografickej oblasti jedinou spoločnosťou, ktorá zahŕňala všetky oblasti elektroenergetiky: výrobu, prenos a distribúciu elektrickej energie. Právny a finančný dohľad bol vykonávaný štátom, ktorý riadil mieru výnosnosti vynaložených investícií a zabezpečoval investovanie finančných prostriedkov do zariadení pre trvalú a spoľahlivú dodávku elektrickej energie, nielen v danom období, ale aj vo výhľade do budúcnosti. Pre zákazníka to síce znamená väčšiu mieru istoty v zabezpečení dodávky elektrickej energie, ale zároveň nesie v cene elektrickej energie investície tejto spoločnosti, či už boli nutné alebo nie.

Snahy a požiadavky veľkých odberateľov, ktorí v záujme využitia výhod plynúcich z rozdielov cien usilovali o možnosť prístupu tiež k iným energetickým podnikom nie len ku svojmu vtedajšiemu dodávateľovi, viedli v priemyselne vyspelých krajinách ku zmene prístupu od vertikálne - integrovaného modelu elektroenergetiky, smerom k otvoreniu konkurencii [1], [2].

V energetickom priemysle na celom svete vplyvom liberalizácie vznikli nové pravidlá ako sú: súťaž, reštrukturalizácia, privatizácia a regulácia. Cieľom týchto reforiem je podporiť hospodársku súťaž za účelom zníženia nákladov, zvýšiť úroveň rozhodovania zákazníkov a konsolidovať verejné financie. Politické nástroje na dosiahnutie týchto cieľov sú: zavedenie hospodárskej súťaže (podporované nariadenie) a podpora účasti súkromného subjektu.

Liberalizáciou trhu sa vytvorilo konkurenčné prostredie aj pri existencii prirodzených monopolov a umožnilo na jednej strane odberateľom elektriny voľbu dodávateľa a na druhej strane ponúklo existenciu rovnocennej súťaže medzi jednotlivými dodávateľmi a tiež postupne vytvorilo prirodzený tlak na zvyšovanie ekonomickej efektívnosti. Trh s elektrinou v SR sa podriaďuje smerovaniu trhu s elektrinou v EÚ. Je tvorený procesmi dodávky elektriny odberateľom na vymedzenom území, prenosom a distribúciou elektriny od jednotlivých dodávateľov. Procesy vyžadujú pripojenie a prístup účastníkov trhu do prenosovej, resp. distribučných sústav [3], [4].

Zákon o energetike 352/2013 Z. z. definuje účastníkov trhu s elektrinou takto:

- výrobca elektriny,
- prevádzkovateľ prenosovej sústavy,
- prevádzkovateľ distribučnej sústavy,
- dodávateľ elektriny,
- odberateľ elektriny,
- organizátor krátkodobého trhu s elektrinou.

### III. ŠTRUKTÚRA CENY ELEKTRINY NA LIBERALIZOVANOM TRHU

Za posledných 10 rokov sa cena elektriny zmenila pokiaľ ide o výšku, ale aj štruktúru. Liberalizácia trhu s elektrinou spôsobila, že finančné toky za jednotlivé druhy elektriny sa rozdelili a koncový odberateľ aj v prípade, že dostáva jedinú faktúru za tzv. združenú dodávku, musí dostať informáciu o jednotlivých zložkách výslednej ceny. Skomplikovanie štruktúry ceny súvisí najmä s politikou poskytovateľov prístupu k distribučnej aj prenosovej sústave. Obaja operátori väčšinu taríf konštruujú ako delenú - tarifa obsahuje stálu platbu, ktorá závisí od času poskytovania prístupu k sústave, tá sa obvykle platí mesačne. Obsahuje vo väčšine prípadov aj výkonovú zložku, ktorá závisí od množstva dodanej elektriny od toho totiž závisí namáhanie zariadení PDS [4].

Nakoľko je cena elektrickej energie závislá na viacerých faktoroch, tak sa môže hranica ceny energie individuálne pohybovať v závislosti od rôznych podmienok. Výška ceny je do určitej miery regulovaná štátnym úradom. Jedným z hlavných dôvodov je prvoradé zabezpečenie prijateľnej ceny pre odberateľov – domácnosti, firemný sektor a priemyselné podniky. Cena elektriny pre koncových odberateľov sa všeobecne skladá z regulovaných a neregulovaných zložiek.

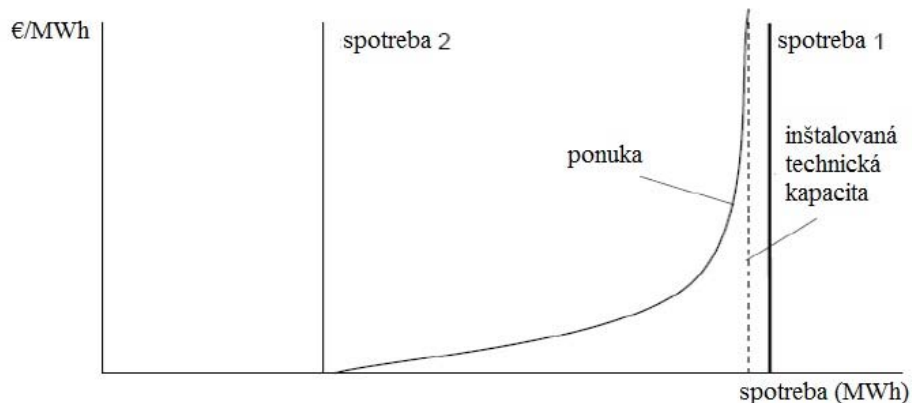
Základom špecifických problémov cenotvorby elektriny je jej neskladovateľnosť. To znamená zabezpečenie rovnováhy medzi ponukou a dopytom, v reálnom čase, vo všetkých uzloch elektrizačnej sústavy.

Zabezpečenie rovnováhy v reálnom čase vyžaduje stálu pohotovosť prenosových ciest (prenosovej sústavy a distribučnej sústavy). Zdroje (dodávatelia) aj spotrebiteľia sú ovplyvňovaní činnosťami ostatných zdrojov a spotrebiteľov aj stavom a prevádzkou prenosových ciest. Pre potreby obchodu s elektrinou, špeciálne pre cenotvorbu bol zavedený pojem spotová cena elektriny (spot price).

Základné úvahy pri formulovaní pojmu spotová cena najmä sú [1]:

- závislosť ceny elektriny v čase,
- závislosť ceny elektriny na umiestnení dodávky či dopytu (spotreby) v rámci elektrizačnej siete,
- stochastičnosť ponuky a dopytu z hľadiska času aj miesta,
- vplyv konkrétnej konfigurácie elektrizačnej sústavy na cenu elektriny.

Na Obr.1 je možné sledovať ponukovú krivku výrobcov pre jednu obchodnú hodinu. Základný princíp, ktorý Obr.1 popisuje je ale platný aj pre jednotlivé produkty a pre dlhšie obdobie.



Obr. 1 Tvorba spotovej ceny elektriny

Za spotovú cenu je považovaná cena, ktorá sa vytvorí na trhu v čase  $d - 1$ , prípadne cena,



ktorá sa vytvorí v najbližší predchádzajúci pracovný deň pred dňom realizácie dodávky. Je to preto, že v nepracovné dni sa neobchoduje, a preto spotovú cenu napríklad na nedeľu a pondelok sa uzatvára už v piatok. Túto cenu možno zároveň považovať za cenu, ktorá je najbližšie k základnej cene. Všeobecne možno totiž povedať, že čím viac sa obchod blíži okamihu spotreby, tým lepšiu majú subjekty obchodu predstavu o skutočných nákladoch výrobcu a predovšetkým presnejšie predikujú spotrebu zákazníkov pre tento okamih [1].

#### IV. BURZY A OBCHODOVANIE S ELEKTRINOU

Obchodovanie s elektrinou v Európe prispelo k nárastu cezhraničného obchodovania, kde voľba dodávateľa vrátane nákupu a predaja elektrickej energie sa neobmedzuje iba na daný štát. Toto cezhraničné obchodovanie však naráža na fyzikálne obmedzenia medzi jednotlivými štátmi alebo regiónmi, pretože prenosové vedenia (vrátane cezhraničných) boli pôvodne určené pre udržanie stability prenosových sústav jednotlivých krajín a zabezpečenie spoľahlivosti dodávok v rámci danej krajiny.

V súčasnosti preto prenosové vedenia čelia obrovským nárokom a požiadavkám na prenos elektriny a predstavujú významné obmedzenia v budovaní jednotného celoeurópskeho trhu s elektrinou. Prvými krokmi, ktoré nediskriminačne a transparentne umožnili prístup k cezhraničnej kapacite prenosových vedení, bolo zavedenie tzv. explicitných aukcií. Explicitnú aukciu možno rozumieť, ako právo na využitie voľného cezhraničného profilu. Tieto aukcie fungujú na základe bilaterálnych dohôd medzi prevádzkovateľmi susediacich prenosových sústav, alebo na báze multilaterálnych dohôd, ak sa vykoná pre viacerých prevádzkovateľov prenosových sústav daného regiónu.

Snaha odstrániť niektoré nevýhody explicitných aukcií a úsilie o dosiahnutie konkurencieschopných trhových cien, dostupnosti elektrickej energie a bezpečnosti dodávok, viedla niektoré štáty k ďalším dohodám o spolupráci prenosových sústav, energetických búr a operátorov trhov s cieľom nájsť nové metódy cezhraničného obchodovania. Výsledkom takýchto dohôd je aplikácia a zavedenie tzv. implicitných aukcií, teda spoločná kúpa elektriny vrátane potrebného cezhraničného profilu v jednom kroku a na jednom mieste [1].

Obchody prebiehajú najmä na burzách. Burzy s elektrinou patria medzi komoditné burzy, sú založené na členskom princípe, čiže obchodovať na jej parkete môžu len registrovaní členovia. Pre registráciu treba splniť náročné kritéria, no aj ostatní účastníci trhu môžu na burze obchodovať a to cez registrovaných členov. Výhodou burzy s elektrickou energiou je, že pre jej členov a ďalších účastníkov poskytuje konkurenčne transparentné prostredie. Môžu ho využívať aj malí dodávateľia elektriny, ktorý nakupujú prostredníctvom člena burzy elektrinu pre svojich odberateľov. V 90-tých rokoch vzniklo v Európe niekoľko búr s elektrinou, z ktorých niektoré obchodujú aj s plynom.

Energetická burza Praha (PXE) je burzou, ktorá vznikla v marci 2007 za účelom obchodovania s energiami s miestom dodania v Českej republike a v roku 2008 a 2009 aj s miestom dodania na Slovensko a v Maďarsku. Pred vznikom burzy PXE, na území Českej republiky neexistoval organizovaný trh, v dôsledku čoho sa obchody uzatvárali na základe bilaterálnych kontraktov. PXE a BCPP (Burza cenných papierov Praha) sú dcérske spoločnosti skupiny PX. Obchodovanie na BCPP prebieha prostredníctvom licencovaných obchodníkov s cennými papiermi, ktorí sú zároveň členmi burzy. Jedná sa predovšetkým o významné banky a maklérske firmy.

Na Energetickej burze Praha sa obchoduje s elektrickou energiou a to výhradne s futures s fyzickým vysporiadaním. Menou, v ktorej sa dané kontrakty dohadujú je euro. Predmetom obchodov na PXE je elektrická energia o hodinovom výkone 1 MW vo všetkých hodinách, všetkých dňoch dohodnutého dodávkového obdobia. Miestom dodania sú elektrizačné sústavy Českej republiky (CZ), Slovenska (SK) a Maďarska (HU), kde doprava elektriny nie je v cene kontraktu zahrnutá. Kontrakty sa delia podľa dĺžky dodania na mesačné, štvrťročné a ročné. Na základe tohto členenia obchoduje PXE so šiestimi typmi produktov, a to mesačnými, štvrťročnými a ročnými BASE LOAD a PEAK LOAD [4], [5], [6].

#### V. ANALÝZA CENOVÉHO VÝVOJA NA TRHU S ELEKTRINOU

Matematické metódy obchodovania ponúkajú objektívny pohľad na cenovú aktivitu. Pomáhajú vytvoriť si názor na smer a pohyb ceny v čase, zmenšiť strach a vyhnúť sa častému obchodovaniu. Okrem toho tieto metódy dokážu určiť novo začatý smer v krátkom čase po zmene smeru cenového pohybu. Pri obchodovaní je možné sa stretnúť s dvoma typmi trhu. Ide o Trending alebo Ranging market (Trendujúci alebo stagnujúci trh).

Ranging Market (Choppy market) je trh, na ktorom výraznejšie neprevláda žiaden trend. Dochádza len k neustálemu striedaniu nastupujúcich rastúcich a klesajúcich trendov, čo sa v konečnom dôsledku prejavuje ich negáciou a bočným pohybom trhu. Opodstatnené je použitie protitrendových indikátorov ako je napr. RSI.

Trending Market je trh, ktorý má jednoznačný smer – či už rastúci alebo klesajúci. Prevládajúci trend umožňuje profitovať z jednoznačného pohybu trhu bez ohľadu nato, či trh rastie alebo klesá. Na trending trhu je vhodné nasadenie Trend- followingových indikátorov a stratégií, akým je napríklad MACD [6].

Nástroje používané v matematických metódach obchodovania sú kľzavé priemery a oscilátory. Hoci existuje nespočetné množstvo matematických indikátorov, tieto sa najviac používajú v praxi [6]:

- Moving Average Convergence-Divergence (MACD - divergencie a konvergencie kľzavých priemerov),
- Index RSI (Relative Strength Index - index relatívnej sily),
- Momentum Indikátor (Rate of change),
- Average directional index (ADX).

Tvorcom MACD je Gerald Appel. MACD radíme k momentum oscilátorom, ktoré vyjadrujú hybnosť, silu momentálneho trendu, pričom oscilujú okolo neutrálnej úrovne oboma smermi. MACD je zložené z viacerých kľzavých priemerov, konkrétne z dvoch alebo troch. Ich zvyčajné nastavenie je 9, 12 a 26 periód. Uvedené periódy sú len orientačné a každý obchodník si ich môže nastaviť podľa vlastných preferencií [6].

Kľzavý priemer – MA (Moving Averages) je v technickej analýze s obľubou používaný najmä kvôli jednoduchosti svojho výpočtu a možnosti kombinovať viacero kľzavých priemerov navzájom, ako to využíva metóda MACD. Často používaný je Exponenciálny kľzavý priemer EMA, pretože dáva vyššiu váhu posledným dňom a ich hodnoty tak majú pri výpočte vyššiu dôležitosť, ako najvzdialenejšie dni.

Pri nastavení hodnôt 26 a 12 periód, vzdialenosť medzi 26 EMA a 12 EMA bude reprezentovať vzdialenosť od MACD po hodnotu nula na indikátore. Čím ďalej ide línia MACD od hodnoty nula tým širšia je medzera medzi 12 EMA a 26 EMA na grafe. Čím bližšie je MACD línia k nule tým bližšie sú aj 12 a 26 EMA. MACD teda meria vzdialenosť medzi MACD čiarou a spúšťacou líniou [6].

Je trendovým indikátorom a preto dáva veľmi dobré signály v trendovej časti trhu. Naopak sa neodporúča používať ho v prípade, že trh sa dostal do netrendovej oblasti. Existuje mnoho spôsobov využitia MACD a obchodníci si ho neustále ďalej prispôbujú vlastným požiadavkám. Jedným z najjednoduchších využití MACD je sledovanie prerazenia MACD histogramu a trigger line. MACD sa počíta z dlhšej periódy napr. MA 26 a trigger line sa počíta z kratšej periódy, ktorou je napr. EMA 12.

Na nasledovnom Obr. 2 sú zachytené MACD hodnoty vo forme histogramu so svojou Trigger line (červenou).

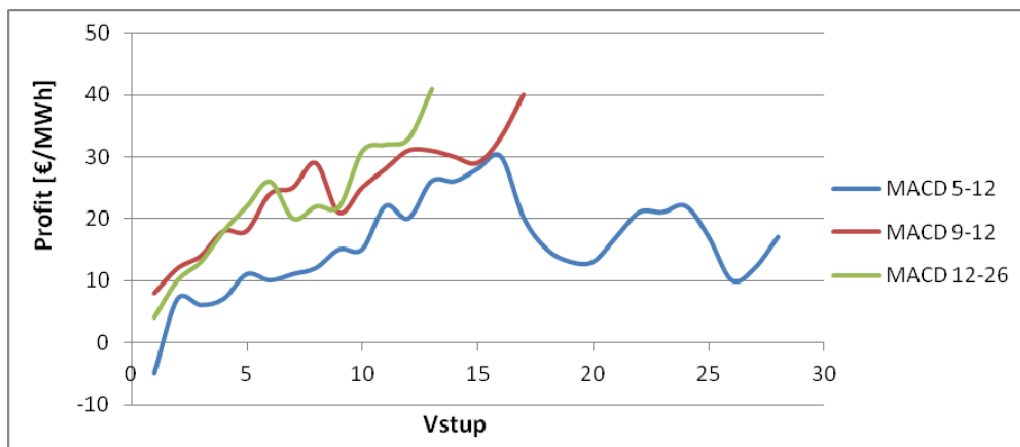


Obr. 2 Obchodovanie pomocou metódy MACD

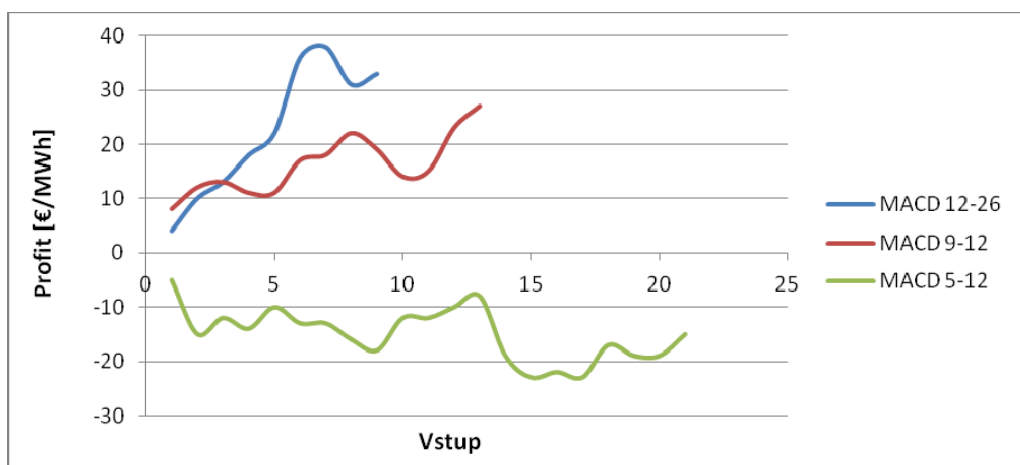
Ak Histogram narastá, znamená to silnejúce Momentum a rastúci rozdiel medzi MACD a Trigger line. Ak Histogram začne klesať, znamená to, že MACD Momentum slabne a je možné vopred predpovedať kríženie MACD – Trigger line a predstihnúť trh s okamihom nákupu/predaja.

Pri využití tejto metódy prekríženia MACD – Trigger line, kde prekrížením pod nulovou osou sa získava signál na nákup a prekrížením nad nulovou osou sa získava signál na predaj.

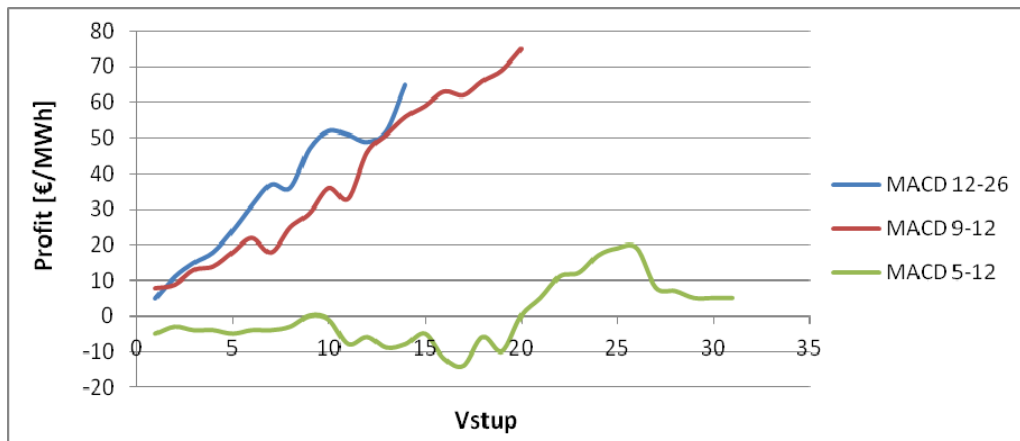
Použitím tejto metódy pri obchodovaní sa dosiahol profit s periódami 5 – 12, 9 – 12 a 12 – 26 dní, ktoré sú porovnané na Obr.3 pre slovenský trh, na Obr.4 pre český trh a na Obr.5 pre maďarský trh.



Obr. 3 Krivky profitu s metódou MACD pri rôznych periódach výpočtu, pre slovenský trh



Obr. 4 Krivky profitu s metódou MACD pri rôznych periódach výpočtu, pre český trh



Obr. 5 Krivky profitu s metódou MACD pri rôznych periódach výpočtu, pre maďarský trh

Krivky profitu zobrazujú výsledný profit po každom signály pre nákup alebo predaj. Jedným vstupom sa rozumie vstup do pozície a zároveň výstup z pozície. Výsledný profit je rozdiel

ceny vstupu a výstupu. Ak bol obchod ziskový a mechanizmus predikoval smer správne, krivka profitov rastie.

Z výsledkov je možné vidieť, že mechanizmus MACD je najviac ziskový s nastavením periódy 12-26. Na všetkých troch trhoch bol tento mechanizmus za sledované obdobie ziskový. Rovnako aj nastavenie periód MACD indikátora na hodnoty 9-12 bol ziskový na všetkých troch trhoch. Najmenej ziskový, resp. v prípade českého trhu stratový bol mechanizmus MACD s periódami 5-12.

## VI. ZÁVER

Nástupom liberalizácie sa zvýšila konkurencieschopnosť subjektov, ktorí dodávajú elektrickú energiu. Vzhľadom na to, že sa s elektrinou obchoduje ako s bežnými komoditami, upriamuje sa pozornosť na jej cenu a možný odhad vývoja ceny. Jedine koncovou cenou pre spotrebiteľa je možné vyhrať konkurenčný boj medzi dodávateľmi elektrickej energie. Dodávatelia nakupujú elektrickú energiu na burzách a pozorne sledujú vývoj cien elektrickej energie. Tento článok pojednáva o aplikácii mechanizmov predikcie ceny elektriny, ktorý je tvorený technickými indikátormi. Technické indikátory sú vypočítané na základe historických cien elektriny. Všetky mechanizmy popísané v tomto príspevku boli aplikované na obdobie od 4.1.2010 do 21.3.2014.

Z výsledkov mechanizmov je možné vyvodit' záver, že technické indikátory sú pri správnom nastavení nápomocné pri predikcii cien elektrickej energie. Avšak rôzne mechanizmy môžu ukázať rôzne signály pre vstup a výstup. Je preto možné predpokladať, že ich kombináciou je možné filtrovať poklesy na krivkách profitov.

## POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva SR a Slovenskej akadémie vied projektom VEGA 1/0388/13.

## LITERATÚRA

- [1] CHEMIŠINEC, MARVAN, et. al.: Obchod s elektrinou. Praha: CONTE spol. s r.o. 2010. 202 s. ISBN 978-80-254-6695-7.
- [2] TÓTH, Roman: Analýza trhu s elektrinou: Diplomová práca. Košice: TU FEI, 2014. 75 s.
- [3] Ministerstvo hospodárstva Slovenskej republiky: Návrh energetickej politiky Slovenskej republiky [online]. Bratislava: MHSR, 2013. Aktualizované 05.2013 [08-09-2013]. Dostupné na internete: <[www.economy.gov.sk/navrh-energetickej-politiky-sr](http://www.economy.gov.sk/navrh-energetickej-politiky-sr)>.
- [4] JANÍČEK, František et. al.: Model trhu s elektrinou. Bratislava: Renesans, s.r.o., 2009. 173 s. ISBN 978-80-89402-03-9.)
- [5] J. Dědič, "Burza cenných papírů a komoditní burza", Praha : Prospektrum, 1992. 295 s., ISBN 80-85431-62-9.
- [6] O. Rejnuš, "Soudobý význam komodit při vztváření nových druhů finančních instrumentů (Strukturovaných produktů)", .In Sborník příspěvků ze 7. mezinárodní vědecké konference – Nové obchodovatelné komodity a jejich využití. 1 vyd. Brno: Masarykova univerzita, 2007, s.35-41, ISBN 978-80-210-4432-6.
- [7] C. Andrieu, M.Fontela, B. Enacheanu, H. Pham, B. Raison, Distributed Network Architectures, European project CRISP, Deliverable D 1.7, 30 August 2005.
- [8] Etrading.sk: Trending vs. Ranging Market [online]. Bratislava: Etrading.sk: 2013. [21-03-2014]. Dostupné na internete: <<http://www.etrading.sk/sk/obchodovanie-na-burze/56/132-trending-vs-ranging-market>>.
- [9] Safecap Investments Ltd.: Matematické indikátory v obchodovaní spoločností [online]. Bratislava. 2014 [2014-04-18]. Dostupné na internete: <<http://www.markets.com/sk/education/technical-analysis/math-indicators.html>>.
- [10] Etrading.sk: MACD – MA Convergence Divergence [online]. Bratislava: Etrading.sk: 2013. [22-03-2014]. Dostupné na internete: <<http://www.etrading.sk/sk/technicka-analyza/44/98-macd-moving-average-convergence-divergence>>.
- [11] Forex-system.: Indikátor MACD (Moving average convergence/divergence) Forex indikátor Divergence [online]. Bratislava: Forex-system: 2013. [22-03-2014]. Dostupné na internete < <http://www.forex-system.sk/news/macd-moving-average-convergence-divergence-forex-indikator/>>.

# Meranie vlastností nelineárnych dielektrík

<sup>1</sup>Radovan MALIAR <sup>2</sup>Ján TKÁČ,

<sup>1</sup> Katedra Elektroenergetiky, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

<sup>2</sup> Katedra Elektroenergetiky, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

<sup>1</sup>jan.tkac@tuke.sk, <sup>2</sup>radovan.maliar@student.tuke.sk

**Abstrakt** — V elektrotechnickom a zvlášť v oblasti výkonovej elektroniky sa často používajú keramické kondenzátory využívajúce nelineárne dielektriká. Tieto materiály majú svoje výhody aj nevýhody. Základnými skúmanými, ale aj používanými materiálmi, ktoré sa používajú vo funkcii nelineárnych dielektrík sú BaTiO<sub>3</sub> a SrTiO<sub>3</sub>. Merania poukázali na skutočnosť, že ich nelinearita veľmi nepriaznivo ovplyvňuje vlastnosti kondenzátorov už pri bežných pracovných teplotách. Tieto kondenzátory sa používajú aj v invertoroch pre solárnu techniku, kde môžu negatívne ovplyvňovať ich životnosť v dôsledku namáhania spôsobujúceho starnutie.

**Kľúčové slová** — nelineárne dielektriká, feroelektriká, curieho teplota, permitivita

## I. ÚVOD

Nelineárne dielektrické materiály sú známe a používané už od začiatku päťdesiatych rokov 20. storočia pre ich výborné materiálové parametre a dostupnú cenu. V oblasti fyziky nelineárnych dielektrických materiálov je v posledných rokoch zaznamenávaný významný pokrok. V súčasnej dobe tieto materiály nachádzajú významné uplatnenie v najrôznejších technických oboroch, preto je ich štúdiu venovaná významná pozornosť. Na kvalitu dielektrických materiálov sú v súčasnosti kladené stále vyššie nároky, či už sú to dielektriká pre kondenzátory, kde je snahou vytvoriť kondenzátory s čo najvyššou hodnotou kapacity pri čo najnižších rozmeroch, ktorá sa v závislosti od času a teploty nemení, so zachovaním cenovej dostupnosti. Pre zabezpečenie stabilných parametrov pri prevádzkových teplotách sú vyvíjané materiály s vysokou hodnotou Curieho teploty.

## II. VLASTNOSTI NELINEÁRNYCH DIELEKTRÍK

Pre nelineárne dielektriká je charakteristická zmena vlastností vo forme teplotnej závislosti. Feroelektrické látky sú charakterizované určitou teplotou  $T_c$ , ktorá je nazývaná Curieho teplota, alebo teplota feroelektrického fázového prechodu. Táto teplota oddeľuje feroelektrický (polárny stav) od stavu paraelektrického (nepolárneho), v ktorom látka nie je spontánne polarizovaná. Oba stavy sa od seba líšia kryštalografickou symetriou. U feroelektrických látok dôjde k zmenám v usporiadaní nábojov. Tento stav potom vedie ku strate spontánnej polarizácie a k poklesu relatívnej permitivity. Takýto pokles je vyjadrený Curie-Weissovým zákonom:

$$\varepsilon' = \frac{C_{cw}}{T - T_c}$$

kde  $C_{cw}$  je Curie-Weissová konštanta,  $T$  je teplota a  $T_c$  je Curieho teplota. Pri teplotách blízkyh  $T_c$  sú tieto látky veľmi poddajné voči pôsobeniu elektrického poľa, z tohto dôvodu je zrejme, že v blízkosti  $T_c$  je možné pozorovať zvláštnu teplotnú závislosť materiálových vlastností, nazývanú diferenciálna permitivita. Paraelektrický stav existuje obvykle pri vyšších teplotách než feroelektrický stav. Feroelektriká pri teplotách blízkyh Curieho bodu sú silne

dielektricky nelineárne a majú vysokú hodnotu relatívnej permitivity. Je však známych niekoľko výnimiek, kedy má látka dva fázové prechody, medzi ktorými existuje feroelektrický stav [3].

### III. MATERIÁLY NELINEÁRNYCH DIELEKTRÍK

Základnými skúmanými, ale aj používanými materiálmi, ktoré sa používajú vo funkcii nelineárnych dielektrík sú  $\text{BaTiO}_3$  a  $\text{SrTiO}_3$ .

Medzi nové materiály z hľadiska výskumu nelineárnych dielektrík patria materiály s chemickým označením  $\text{Sr}_2\text{Nb}_2\text{O}_7$  a  $\text{Ca}_2\text{Nb}_2\text{O}_7$ . Sú vhodnými kandidátmi pre vysoko-teplotné piezoelektrické aplikácie z dôvodu ich vysokej hodnoty Curieho teploty.  $\text{Sr}_2\text{Nb}_2\text{O}_7$  (SNO) a  $\text{Ca}_2\text{Nb}_2\text{O}_7$  (CNO) a iné piezoelektrické materiály so štruktúrou  $\text{A}_2\text{B}_2\text{O}_7$  patria do perovskitovej vrstvenej štruktúry materiálov označovanej PLS, kde všeobecný vzorec má formu  $\text{A}_n\text{B}_n\text{O}_{3n+2}$  ( $n = 4$  pre  $\text{A}_2\text{B}_2\text{O}_7$ ). Skúmaním  $\text{Sr}_2\text{Nb}_2\text{O}_7$  bola nameraná Curieho teplota  $T_c = 1327 \pm 5^\circ\text{C}$  a pre  $\text{Ca}_2\text{Nb}_2\text{O}_7$  bola nameraná hodnota  $T_c = 1525^\circ\text{C}$  [2].

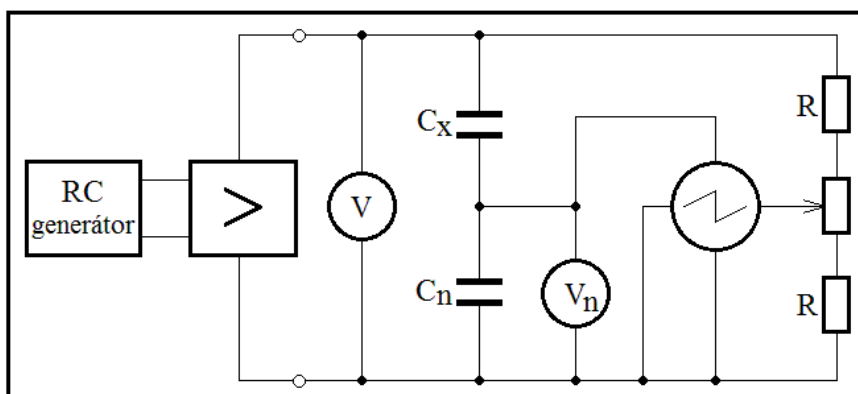
Jednofázové husté  $\text{Sr}_2\text{Nb}_2\text{O}_7$  a  $\text{Ca}_2\text{Nb}_2\text{O}_7$  keramiky s orientovanou štruktúrou mikroštruktúry sa vyrábajú plazmovým spekaním použitím dvojstupňovej metódy. Plazmové spekanie využíva pulzný jednosmerný prúd a jednoosí tlak na konsolidáciu prášku. V prvom stupni je prášok SNO spekaný pri teplote  $1200^\circ\text{C}$  a tlaku  $82\text{ MPa}$  po dobu 3 minút. V druhom stupni sa vzorka deformuje cez väčšiu trysku a je spekaná pri teplote  $1350^\circ\text{C}$  a tlaku  $80\text{ MPa}$  po dobu 5 minút. Spekacie podmienky pre CNO sú takmer rovnaké, s výnimkou teploty spekania pre druhý stupeň, ktorej hodnota bola  $1400^\circ\text{C}$ . Spekané vzorky SNO a CNO sú následne žihané pri teplote  $1200^\circ\text{C}$  po dobu 15 hodín na odstránenie znečistenia uhlíkom a iných nečistôt [2].

V systéme  $\text{CaO-SrO-Bi}_2\text{O}_3\text{-Nb}_2\text{O}_5\text{-Ta}_2\text{O}_5$  existuje niekoľko fáz (zmesí oxidov), ktoré sú predmetom záujmu základného a aplikovaného výskumu. Relatívne nový a prakticky najvýznamnejší z nich je  $\text{SrBi}_2\text{Ta}_2\text{O}_9$  (SBT), ktorý je v súčasnej dobe komerčne využívaný ako materiál pre výrobu FRAM. SBT tak predstavuje alternatívu k bežne používanému materiálu  $\text{Pb(Zr, Ti)O}_3$ , pričom okrem nižšej toxicity má aj ďalšie výhody, ako kratšiu prístupovú dobu k informáciám a dlhšiu životnosť. SBT patrí medzi oxidy s vrstvenatou štruktúrou typu  $[\text{Bi}_2\text{O}_2]^{2+}[\text{A}_{m-1}\text{B}_m\text{O}_{3m-1}]^{2-}$ . Curieho teplota  $T_c$  prechodu z feroelektrického do paraelektrického stavu bola pre tieto vzorky stanovená v rozmedzí  $270$  až  $335^\circ\text{C}$ . Curieho teplotu rovnako ako ďalšie elektrické parametre (relatívnu permitivitu, remanentnú polarizáciu, koercitívnu silu) SBT je možné ovplyvniť (riadiť) čiastočnou alebo úplnou substitúciou tantalu a nióbu, resp. stroncia a vápnika (napr. Nb resp. Ca zvyšuje Curieho teplotu  $T_c$ ) [1].

### IV. EXPERIMENT

Pomocou osciloskopickej metódy bola určovaná kapacita a permitivita skúmaných vzoriek v širokom rozsahu hodnôt intenzít elektrického poľa a teplôt. Výstupom z osciloskopu je hysterézná slučka. Základné zapojenie obvodu je zobrazené na Obr. 1.

Zdrojom je RC generátor, ktorý dodáva napätie s požadovanou frekvenciou do zosilňovača. Napätie je ďalej zosilnené a privedené na meranú vzorku s kapacitou  $C_x$ . Do série so vzorkou  $C_x$  je zapojený normálový kondenzátor. Napätie z kondenzátora  $C_n$  odpovedajúce náboju na vzorke je privedené na vertikálny zosilňovač osciloskopu. Na horizontálny zosilňovač osciloskopu je privedené vstupné napätie, teda napätie na vzorke upravené odporovým deličom.



Obr. 1 Základné zapojenie pre meranie osciloskopickou metódou

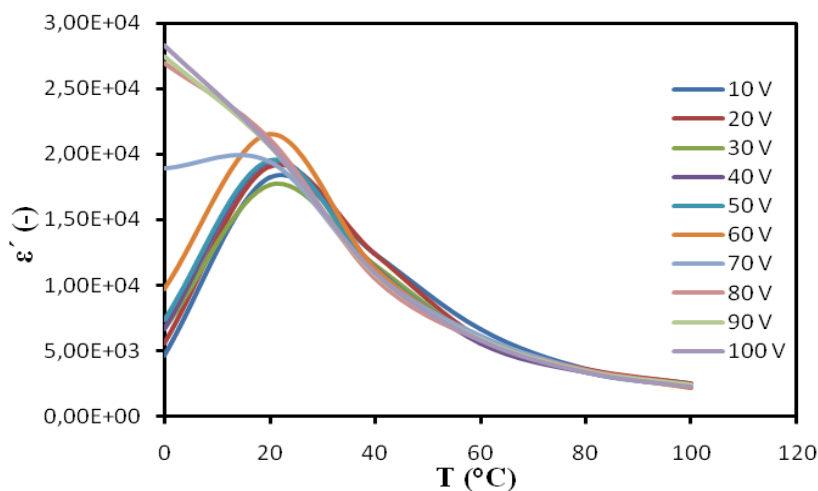
Touto metódou merania boli merané vlastnosti vzoriek z Tabuľky 1 a Tabuľky 2. Vzorky boli uložené v olejovej nádobe a dostatočne vytemperované a pri teplotách 0, 20, 40, 60, 80 a 100 °C. Na vstupné svorky bolo privádzané elektrické napätie  $U_x$  (max. hodnoty) o veľkosti 10 až 100 V s krokom 10 V. Rozmery vzoriek a ich vypočítané geometrické kapacity sú uvedené v Tabuľke 1 a Tabuľke 2.

Tabuľka 1  
 Parametre vzoriek 1, 2, 3 a 4

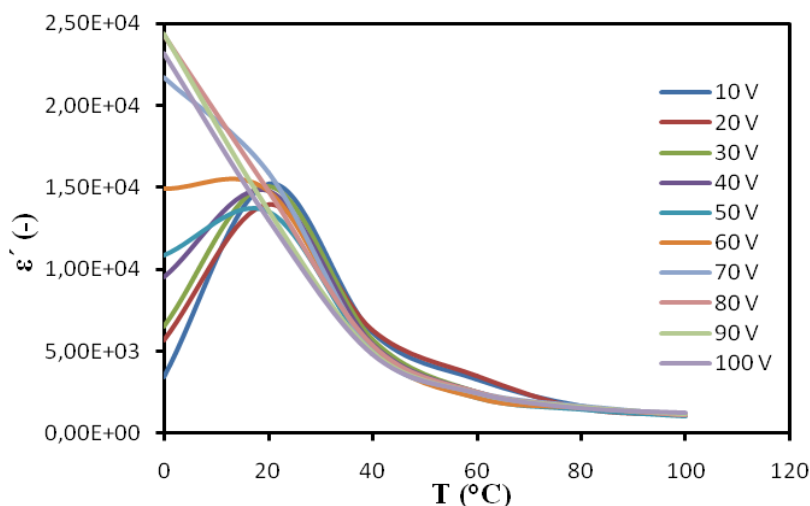
Vzorky	Kapacita [nF]	Napät'ová hladina [kV]	h [mm]	d [mm]	$C_0$ [F]
Vzorka 1	10,3	3	1,8	12,5	$6,036 \cdot 10^{-13}$
Vzorka 2	0,471	3	1,8	5,1	$1,005 \cdot 10^{-13}$
Vzorka 3	1	22	1,8	13,8	$7,357 \cdot 10^{-13}$
Vzorka 4	0,36	6	2,3	30,5	$2,813 \cdot 10^{-12}$

Tabuľka 2  
 Parametre vzoriek 10, 11, 12 a 13

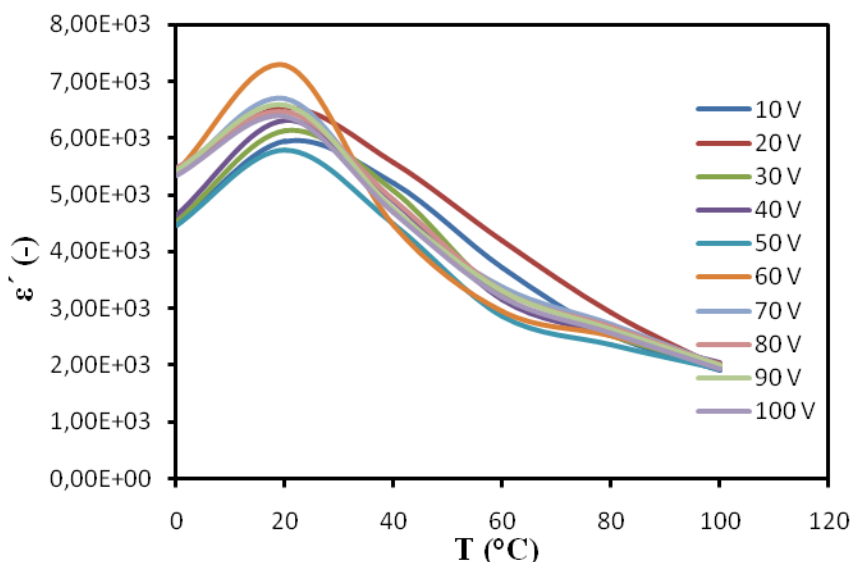
Vzorky	Kapacita [nF]	Napät'ová hladina [kV]	h [mm]	d [mm]	$C_0$ [F]
Vzorka 10	10	0,23	1,45	15,5	$1,152 \cdot 10^{-12}$
Vzorka 11	5,5	10	4,8	25,5	$9,42 \cdot 10^{-13}$
Vzorka 12	4,7	10	4,1	25,5	$1,103 \cdot 10^{-12}$
Vzorka 13	21	100	0,5	20	$5,563 \cdot 10^{-12}$



Obr. 2 Závislosť permitivity vzorky 1 od teploty



Obr.3 Závislosť permitivity vzorky 10 od teploty



Obr.4 Závislosť permitivity vzorky 11 od teploty

Na základe nameraných a vypočítaných hodnôt a grafických závislostí boli jednotlivé vzorky rozdelené do dvoch skupín. Prvú skupinu tvoria vzorky 1, 10 a 11, pri ktorých z nameraných výsledkov bolo možné určiť ich Curieho teplotu. V meranom teplotnom rozmedzí vykazovala ich permitivita maximálne hodnoty pri pôsobení napätia. Druhú skupinu tvoria ostatné vzorky, pri ktorých z nameraných výsledkov nebolo možné určiť ich Curieho teplotu. Závislosť ich permitivity od napätia a teploty mierne stúpala, klesala, alebo kolísala.

Na Obr. 2 je zobrazená grafická závislosť permitivity vzorky 1 od napätia ktorej  $T_c$  je  $17\text{ }^\circ\text{C}$ . Z grafickej závislosti vyplýva, že pri teplote  $0\text{ }^\circ\text{C}$ , kedy je vzorka 1 vo feroelektrickom stave, a pri malých hodnotách striedavého napätia (10 V), nedochádza v sústave domén k pohybu doménových stien, takže úroveň polarizácie je nízka a tejto úrovni odpovedá počiatková relatívna permitivita. Pri zvyšovaní intenzity striedavého napätia až do hodnoty 100 V sa polarizácia dielektrika zvyšuje a rovnaký charakter zmeny vykazuje aj permitivita. Pri zvýšení teploty na  $20\text{ }^\circ\text{C}$  vzorka 1 prechádza z feroelektrického do paraelektrického stavu, kedy zaniká jej spontánna polarizácia. Z Obr. 2, na ktorom je zobrazená závislosť permitivity vzorky 1 od teploty je zrejmé, že s ďalším zvyšovaním teploty dochádza k poklesu permitivity.

Na Obr. 3 je zobrazená grafická závislosť permitivity vzorky 10 od napätia, ktorej  $T_c$  je  $16\text{ }^\circ\text{C}$ . Z grafických závislostí zobrazených na Obr. 3 vyplýva, že vzorka č. 10 má v závislosti od veľkosti napätia a teploty rovnaký priebeh charakteristík ako vzorka 1.

Na Obr. 4 je zobrazená grafická závislosť permitivity vzorky 11 od napätia, ktorej  $T_c$  je



23,5 °C. Z grafickej závislosti vyplýva že hodnota permitivity sa v závislosti od veľkosti striedavého napätia mení iba minimálne, čo je spôsobené tým, že je navrhnutá pre vyššiu napäťovú hladinu. Avšak z grafickej závislosti vyplýva, že permitivita vzorky 11 so stúpajúcou teplotou rastie až do hodnoty  $T_c$ , kde dosahuje svoju maximálnu hodnotu a s ďalším zvyšovaním teploty sa dostáva do paraelektrického stavu a hodnota jej permitivity klesá.

#### V. ZÁVER

Z nameraných výsledkov vyplýva, že Curieho teplota vzoriek 1, 10 a 11 sa nachádza v oblasti ich bežných pracovných teplôt. Zmena dosahovala niekoľkonásobok, ale boli namerané aj rádove zmeny permitivity a kapacity. Použitie týchto vzoriek dielektrík ako kondenzátorov v elektrotechnike je preto nevhodné, nakoľko sa v oblasti Curieho teploty správajú veľmi nestabilne a nelineárne. Kapacita takýchto kondenzátorov je teda veľmi závislá od zmeny teplôt. Pre vzorku 12 a ostatné merané vzorky z výsledkov merania vyplýva, že Curieho teplota týchto vzoriek sa nachádza mimo oblasti ich pracovných teplôt. V oblasti ich pracovných teplôt sa teda využíva ich lineárna a stabilná časť charakteristiky.

V súčasnosti sú vyvíjané nelineárne dielektriká, na ktoré sú kladené podmienky zvyšovania ich elektrickej pevnosti a najmä zvyšovania ich hodnôt Curieho teploty, aby sa v oblasti ich pracovných teplôt využívala najmä ich lineárna časť charakteristiky. Takýmito materiálmi sú napríklad  $\text{SrBi}_2\text{Ta}_2\text{O}_9$  používaný na výrobu FRAM pamätí,  $\text{Sr}_2\text{Nb}_2\text{O}_7$  a  $\text{Ca}_2\text{Nb}_2\text{O}_7$  používané v rôznych piezoelektrických aplikáciách.

#### POĎAKOVANIE

Táto práca vznikla za podpory vedeckej grantovej agentúry ministerstva školstva SR a Slovenskej akadémie vied v rámci projektu VEGA 1/0487/12 a APVV-0185-10.

#### LITERATÚRA

- [1] LEITNER, J. Thermodynamic properties of mixed oxides in the CaO-SrO-Bi<sub>2</sub>O<sub>3</sub>-Nb<sub>2</sub>O<sub>5</sub>-Ta<sub>2</sub>O<sub>5</sub> system | *Termodynamické vlastnosti smesných oxidů v systému CaO-SrO-Bi<sub>2</sub>O<sub>3</sub>-Nb<sub>2</sub>O<sub>5</sub>-Ta<sub>2</sub>O<sub>5</sub>* [online]. 2010 [cit. 2014-03-17].
- [2] Ning, H., Yan, H. and Reece, M. J. (2010), Piezoelectric Strontium Niobate and Calcium Niobate Ceramics with Super-High Curie Points. *Journal of the American Ceramic Society*, 93: 1409–1413
- [3] BŘEZINA, Bohuslav. *Feroelektrika*. 1. vyd. n. p. Praha: Academia, 1973. 200 s. ISBN 21-018-73.

# Application for 3D morphing

<sup>1</sup>Branislav SOBOTA, <sup>2</sup>Tawfik MUDARRI

<sup>1</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>branislav.sobota@tuke.sk, <sup>2</sup>tawfik.mudarri@tuke.sk

**Abstract** — This paper deals with the morphing process between source and target objects. Paper is divided in to four parts. The first part is explanation of morping process. Second part is goal of our solution. Next part gives a brief of analyze 3D morpfing. The last part is engaged in detail the morphing process. This part shows practical results with morphing process. The result is an application created under the XNA Framework allowing people to create morphing process. In the conclusion is advantage and disadvantage of this solution. This solution is usable for visual simulation of some morphing processes of virtual objects in created virtual-reality system.

**Keywords** — morphing, warping, 3D object, XNA framework, virtual reality

## I. INTRODUCTION

Some virtual objects used in virtual space are changed in time or, for example, after collision with other objects. One solution is use of 3D morphing. Morphing is a special effect used in film or animation, where there is switching from one image/object to another, while there is a metamorphosis of the source object to target object according to predefined rules. Objective is to create such a sequence, which would be interpreted as a continuous transformation of one object to another. In the film industry, the effect is used quite often. Recently, with the development of 3D graphics is also in this area discovered the concept of morphing. The aim is to morphing the 3D objects between them according to some predefined properties. Goal of this work is to create an application with morphing by Ferguson's curve. Classically morphing is usually linear, it is done in a straight line.

## II. ANALYSIS

Objects in the boundary representation are described by vertices and faces, so the model  $M$  is a pair  $(V, F)$ , where  $V$  is the set of vertices, each vertex  $V_i$  is described by its coordinates in  $R^3$ . The set  $F$  is the set of faces, where each face  $F_i$  is described by a triple  $(j, k, l)$ , where  $j, k, l$  are indices of vertices which form the face  $F_i$ . In further text we will denote the model described by the pair  $(V, F)$  as a *mesh* [1].

In the classical setting of mesh morphing we have two input meshes. The source mesh  $M_0 = (V_0, F_0)$  and the target mesh  $M_1 = (V_1, F_1)$ . Note that unlike the approach in tools such as 3DS max and Lightwave, the source and the target mesh can have a different number of vertices and a different topology. The term topology refers here to the connectivity graph of vertices. As the source and the target mesh may have a different number of vertices, the first step is establishing of correspondence. The most convenient element for establishing correspondence is a vertex - so we establish the correspondence of vertices of the source and the target mesh. Due to different number of vertices of the source and the target mesh, we cannot link to one vertex of the source mesh exactly one corresponding vertex of the target mesh. Instead, we assign to one vertex from the source mesh generally some place on the surface of the target mesh (i.e. not necessarily a vertex). This situation is illustrated in the Fig. 1. The green arrows represent the relation of correspondence between a few vertices of the source mesh and places on the surfaces of the target mesh. Note that in this way we established a correspondence only for the vertices of the source mesh, so the same process have to be repeated for the vertices of the target mesh, with that difference that we look for an appropriate place on the surface of the source mesh [2].

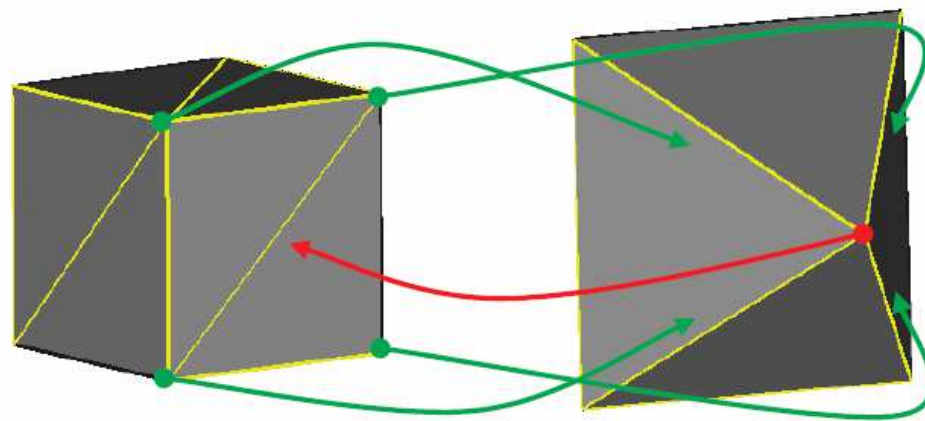


Fig. 1 The correspondence of the vertices of the source mesh with the places of the surface of the target mesh (green arrows) and the correspondence of the vertex of the target mesh with the places on the surface of the source mesh (red arrow).

In this step it is also possible to include some user interaction, usually in the form that user specifies some corresponding features of both input objects, which should remain aligned during the shape transformation. A typical example is morphing between two animals, where we usually want the common features (i.e. legs, head, tail, etc.) to remain aligned throughout the transition.

Once we have established a correspondence, we are still not able to complete morphing by interpolation of corresponding elements, because the correspondence relation is a relation between the vertex of one mesh and the place on the surface of the other mesh. Therefore, the source and the target mesh have to be refined in order to have a vertex-vertex correspondence, i.e. on the positions of corresponding places new vertices have to be inserted. So the second step of the morphing task is a construction of a new representation, which is based on merging of the topologies of the source and the target mesh. This process is also denoted as the topological merging. The new representation is again a mesh, which we will call a *supermesh* in the further text, because it shares a topology of the source and the target mesh. *Supermesh* has one nice property - it can be transformed to the shape of the source mesh as well as to the shape of the target mesh. This situation is shown in the Fig. 2, there are the same objects as on Fig. 1 and the blue dotted lines represent how the topology of the source, and the target mesh respectively, has to be refined in order to be able to transform the *sourcemesh* to the shape of the target mesh and vice-versa [3][4][5].

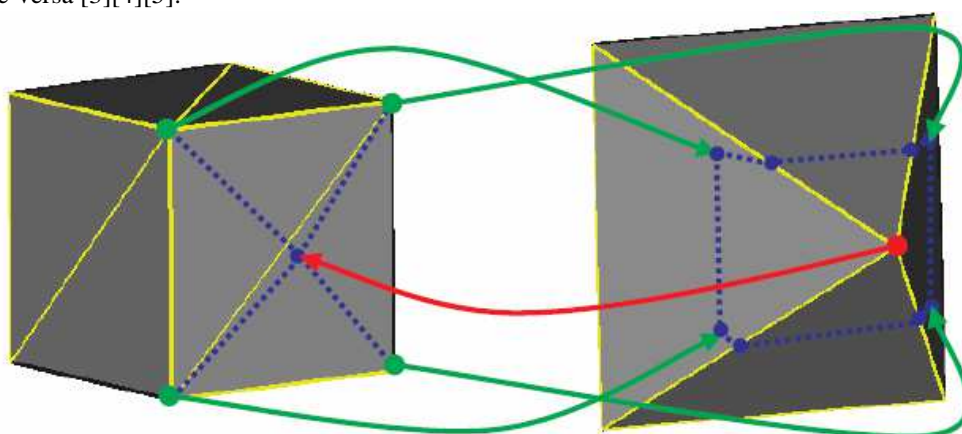


Fig. 2 Refining the topology (blue dotted lines) of the source mesh, and the target mesh respectively, in order to obtain the *supermesh*, which can be transformed to the shape of the source mesh as well as to the shape of the target mesh.

Let us summarize the described steps:

- Establishing of a correspondence of vertices
- Construction of the supermesh, which represents both: the source and the target mesh.
- Interpolation of the corresponding vertices along with surface attributes.

### III. SOLUTION AND RESULTS

3D morphing application Fig. 3 is implemented in C# development environment in MS Visual Studio using 3.2 XNA([6]) framework.

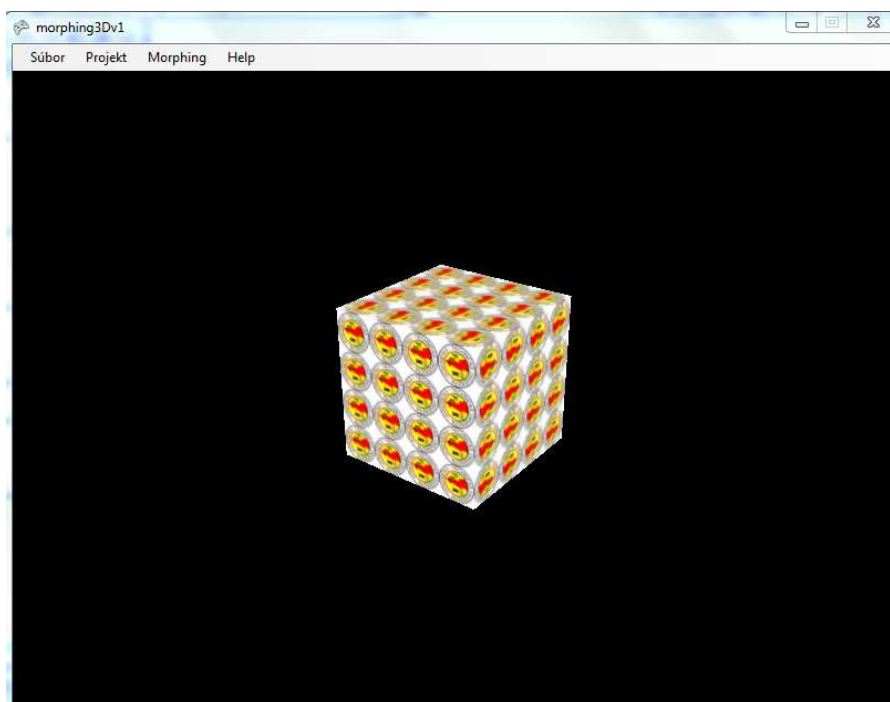


Fig. 3 3D morphing application

#### A. 3D morphing application

The main benefit of the application is that objects among themselves may morphing two different ways: a straight line and a curve (Ferguson's curve). Furthermore, it is possible to set the time course morphing into a linear, i.e. morphing of the initial 3D object to a final 3D object takes place uniformly in time. The second option is to set the non-linear time change where the user has the option to choose the number of periods over the course of morphing and then assign each period of time in seconds. In this way it is possible to create morphing with a very interesting effect. The user has the option of excluding choice of new 3D objects and predefined set of 5 morphing: morphing cube (cube), ball (sphere), pyramid (cone), torus (torus) and monkey object (part Blender). Application to display models in 3 modes:

- Solid
- Wirefire
- Point cloud

During the entire course of morphing is possible to manipulate the object in all directions (X, Y, Z) and rotated about all axes. Morphing using a curve can be achieved very much animation, other than classical morphing straight line (see Fig. 4).



Fig. 3 Example of the morphing by Ferguson's curve - the time course from  $t = 0$  to  $t = 1$

#### IV. CONCLUSION

The advantage of this solution is that the user can choose their own source and target 3D object and set its own course morphing. The user can its own course morphing record the video footage as well as save your project for future work. Limitation of this application is that the source and destination object must have the same number of vertices and surfaces. Another advantage is that the user has the editing mode of vertices on the target object and thus can change the shapes of 3D objects.

Appropriate amendment of this application would be if the application support more 3D formats as well as the addition algorithm, which would be able to convert and then optionally optimize accordingly tops to make them morphing 3D objects, which have different number of vertices.

Next work will be focused to morphing process without limitation of the same number of vertices between source and destination object.

#### ACKNOWLEDGMENT

This work is supported by KEGA grant project No. 054TUKE-4/2013: Application of virtual-reality technologies for handicapped persons education and this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF (25%).

#### REFERENCES

- [1] Kent, J., R., Carlson, W., E., Parent, R., E.: Shape Transformation for Polyhedral Objects, *Computer Graphics*, Vol. 26, Issue 2, 1992, pp. 47-54.
- [2] Gomes, J., Darsa, L., Costa, B., Velho, L.: *Morphing and Warping of Graphical Objects*, Morgan Kaufmann, 1999
- [3] Lazarus, F., Verroust, A.: Three-dimensional metamorphosis: a survey, *The Visual Computer*, Vol. 14, Issue 8 - 9, 1998, pp. 373 - 389.
- [4] Agoston, M. K., 2005. *Computer graphics and geometric modeling: implementation and algorithm*. Springer, 2005. 907 s. ISBN 1852338180.
- [5] Foley, J. D., 1996. *Computer graphics: principles and practice*. Addison-Wesley Professional, 1996. 1175 s. ISBN 0201848406.
- [6] Evangelista, B. P.; Lobao, A. S.; Grootjans, R.; aj.: *Beginning XNA 3.0 Game Programming: From Novice to Professional (Expert's Voice in XNA)*. Apress, 2009, ISBN 978-1430218173, 1 edition.

# Prehľad využitia metód dolovania v dátach vo výrobnom priemysle

Alexandra LUKÁČOVÁ, Veronika KALAVSKÁ, Ján PARALIČ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

alexandra.lukacova@tuke.sk, veronika.kalavska@student.tuke.sk, jan.paralic@tuke.sk

**Abstrakt** — Automatizovaný zber a ukladanie dát vo výrobe predstavuje v dnešnej konkurenčnej dobe základnú potrebu. Na kvalitu finálnych výrobkov sa kladie veľký dôraz, nakoľko priamo ovplyvňuje tak spokojnosť zákazníka ako aj zisk výrobnej firmy. Objavovanie znalostí by v tejto oblasti mohlo pracovníkom zjednodušiť podporu rozhodovania. Hlavným cieľom tejto práce je poskytnúť prehľad problematiky zaoberajúcej sa aplikovaním metód dolovania v dátach z výroby, so zameraním sa na zlepšovanie kvality finálnych výrobkov, a na praktickom príklade z oblasti výroby polyamidových vlákien identifikovať možné smery ďalšieho výskumu.

**Kľúčové slová** — predikcia, klasifikácia, výrobný priemysel, kontrola kvality, predspracovanie, objavovanie znalostí

## I. ÚVOD

V uplynulých rokoch čoraz viac spoločností z priemyselných odvetví zameriava svoju pozornosť na proces dolovania v dátach, pretože vďaka znalostiam môžu spoločnosti lepšie predvídať a zlepšovať výrobné procesy a získať tak konkurenčnú výhodu [1]. Autori práce [2] tvrdia, že kontrola výstupu môže byť vykonaná dvoma spôsobmi, a to buď sledovaním premenných a vykonávaním ich korekcií v reálnom čase alebo tzv. "posthoc" prístupom, kde sa porovnáva konečný výsledok s požadovanými špecifikáciami, analyzujú sa príčiny problémov a nastavujú parametre tak, aby zlepšili budúcu kvalitu výrobkov. V práci využívame druhý prístup, pričom sa pridrižujeme metodológii CRISP-DM (z angl. Cross-Industry Standard Process for Data Mining), bližšie popísanej v [3], [4] a zdôrazňujeme spoluprácu s doménovým expertom (viď Obr. 1). Aplikácie s využitím metód dolovania v dátach vo výrobe môžeme vidieť vo viacerých oblastiach ako sú napr. dizajn produktu, kontrola kvality, rozvrhovanie alebo odhad dodacej lehoty (vrátane výroby) od prijatia objednávky [5]. Podľa [6] sa môžeme na úlohy zamerané na zlepšovanie kvality vo výrobe pozerat' zo štyroch pohľadov, a to:

- popis produktu a identifikovanie premenných, ktoré významne vplývajú na jeho finálnu kvalitu (tieto aktivity sa týkajú prvých fáz CRISP-DM, pochopenie problému a pochopenie dát),
- predikcia kvality, ak je sledovaný atribút numerického charakteru,
- klasifikácia kvality, ak je výstupná premenná nominálneho, ordinálneho alebo binárneho charakteru (napr. defekty),
- optimalizácia parametrov na základe charakteristík zistených z prípadov, ktoré viedli k výbornej kvalite.

V nižšie uvedenom prehľade (kap. II) sa zameriame hlavne na prvé tri z uvedených úloh. Poslednú časť (kap. III) venujeme návrhu možného využitia metód dolovania v dátach v procese výroby polyamidových vlákien.

## II. ANALÝZA SÚČASNÉHO STAVU

Problematika využitia objavovania znalostí z veľkých databáz získaných z priemyslu bola skúmaná vo viacerých článkoch. V tomto prehľade sa zameriavame na použité spôsoby predspracovania dát a rôzne modelovacie techniky pri popise, klasifikácii a predikcii kvality výrobkov.

### A. Predspracovanie dát

Dáta z reálneho sveta sú často neúplné, nekonzistentné a obsahujú šumy. A preto hlavným cieľom fázy predspracovania dát je vyriešiť tieto problémy, a tým zvýšiť kvalitu dát, pretože bez kvalitných dát nie je možné dosiahnuť dobré výsledky dolovania v dátach. Tabuľka 1 poskytuje súhrnný prehľad použitých metód vo fáze predspracovania dát.

Tabuľka 1  
 Súhrnný prehľad riešenia problémov v jednotlivých článkoch vo fáze predspracovanie dát

Problém	Metóda	Článok
chýbajúce hodnoty	nahradenie priemerom vymazanie	výroba kobercov [1] výroba potravín [7]
veľké množstvo dát	redukcia vyhľadanim dôležitých minim a maxim	žiarové zinkovanie ocele [8]
veľké množstvo atribútov	redukcia na základe informačného zisku	výroba kobercov [1]
	redukcia na základe Pearsonovho korelačného koeficientu	výroba potravín [7]
nevyvážená trieda	kombinácia „Tomek“ línií [9] a algoritmu náhodných stromov	veterné turbíny [10]
časový rad obsahuje vlny s vysokou alebo nízkou frekvenciou	rýchla fourierova transformácia	žiarové zinkovanie ocele [8]
časový rad s rušivými hodnotami	filtre využívajúcu hraničnú hodnotu	žiarové zinkovanie ocele [8]
numerické hodnoty	transformácia na nominálny atribút pomocou zadaných intervalov	automobilový priemysel [10]

### B. Popis produktu

V [8] je popísaná prípadová štúdia spoločnosti pôsobiacej v oceľiarskom odvetví, kde aplikovali metódy dolovania v dátach pre zistenie príčin zníženia kvality ocele pri procese zinkovania. Teda cieľom bolo identifikovať vplyv zinkovania a jeho parametrov na príľnavosť zinku, ktorý je jedným z kľúčových faktorov odolnosti pozinkovaného výrobku proti korózií. Táto odolnosť je ovplyvnená viacerými faktormi ako napríklad zložením a teplotou zinkového kúpeľa, hrúbkou povlaku a pod. Nastavenie týchto parametrov je veľmi prácná úloha, nakoľko sa tieto parametre líšia v závislosti od typu a rozmerov spracovávaného kusu ocele. Boli vybrané atribúty, ktoré by mali mať najväčší vplyv na proces zinkovania a tiež cieľový atribút, ktorý mal binárny charakter, a to či hrúbka povlaku zinku bola dobrá alebo nie. Pri procese dolovania v dátach spoločnosť zvolila využitie asociačných pravidiel, a to s minimálnou podporou pravidla 50 %. Získané znalosti vysvetlili niektoré príčiny zníženia kvality oceľových zvitkov a poukázali na tie časti procesu výroby, na ktoré sa bolo potrebné zamerať.

Ďalším príkladom je štúdia [11] spoločnosti pôsobiacej v automobilovom priemysle, kde postupovali podľa známej metodiky CRISP-DM a tiež dali prednosť využitiu asociačných pravidiel (algoritmus Apriori; z dôvodu záujmu nájdenia aj zriedkavých vzorov, bola minimálna podpora nastavená veľmi nízko). Vstupnými atribútmi boli rozmery jednotlivých vyrobených častí automobilov. V závere pre potvrdenie významnosti pravidiel využili  $\chi^2$  test nezávislosti.

### C. Klasifikácia kvality

Existuje široký výber algoritmov vhodných pre ich aplikovanie v klasifikačných úlohách. Práca [12] radí medzi najpopulárnejšie CART [13], ID3 [14] a jeho potomkov C4.5 [15] a C5.0 [16].

Štúdia [1] spracováva prípad tureckej spoločnosti, ktorá sa zaoberá priemyselnou výrobou kobercov. Hlavným cieľom bolo definovať príčiny porúch výrobných zariadení. Použili sa dáta v rozmedzí jeden a pol roka, kde každý záznam predstavoval jednu poruchu zariadenia, ktorá bola popísaná atribútmi ako dátum, číslo zariadenia, typ poruchy, ID poruchy, dôvod poruchy a trvanie poruchy. Cieľovým atribútom bol v tomto prípade binárny atribút trvanie poruchy, kde dané trvanie poruchy bolo klasifikované buď ako „normálne“ alebo „abnormálne“. Chýbajúce hodnoty nahrádzali priemernou hodnotou. Modelovanie bolo vykonané prostredníctvom rozhodovacieho stromu C4.5, z ktorého bolo vygenerovaných približne sto pravidiel. Presnosť modelu bola 72,81%.

Cieľom článku [17] bolo porovnať navrhnutý BOW algoritmus (z angl. Breadth-Oblivious-Wrapper) s algoritmom C4.5. Použité dáta pochádzajú z potravinárskeho priemyslu, konkrétne z oblasti výroby mliečnych výrobkov - výroby tvarohu. Je prirodzené, že sa tvaroh počas svojej doby trvanlivosti a najmä k jej koncu stáva kyslejší, a to často predstavuje problém pre spotrebiteľov, pretože sa domnievajú, že ide o pokazený produkt. Úlohou bolo určiť dávky,



ktoré sa s vysokou pravdepodobnosťou stanú kyslé na konci doby trvanlivosti. Dáta obsahovali približne 800 záznamov, pričom každý záznam mal 70 vstupných atribútov reprezentujúcich rôzne výrobné premenné, ktoré spadali do skupín ako teplota, trvanie, suroviny, prístroj a pod. Cieľový atribút predstavoval hodnotu pH po dvoch týždňoch, teda tesne pred koncom doby trvanlivosti. Tento atribút obsahoval dve hodnoty, a to „chutný“ (pH v rozmedzí 4,9 – 5,3) a „kyslý“ (pH pod 4,9). Na základe výsledkov zistili, že algoritmus BOW (presnosť 85,92%) je účinnejší ako algoritmus C4.5 (presnosť 77,52%).

Inovatívny prístup preukazuje práca [7], kde autori navrhujú RapidMiner rámec pre analýzu výrobných dát v klaude. Experimenty vykonávali na základe dát (600 záznamov, 28 atribútov) z firmy, ktorá sa zaoberala výrobou jedla, pri modelovaní využili algoritmus C4.5 (výstupom bol atribút, ktorý značil, či pomer medzi celkovou výslednou hmotnosťou a celkovou vstupnou hmotnosťou každej výrobnej šarže bol menší ako 90%) a na vyhodnotenie bola použitá presnosť. Výstupom sú pravidlá, ktoré naznačujú, ktoré parametre by mali byť kontrolované a upravené. Dosiahnutá presnosť výsledného modelu predstavovala 95,3%. Navyše, rámec pre analýzu dát bol implementovaný ako plug-in pre RapidMiner, čím poskytol väčší prístup k analýze dát i pracovníkom výroby.

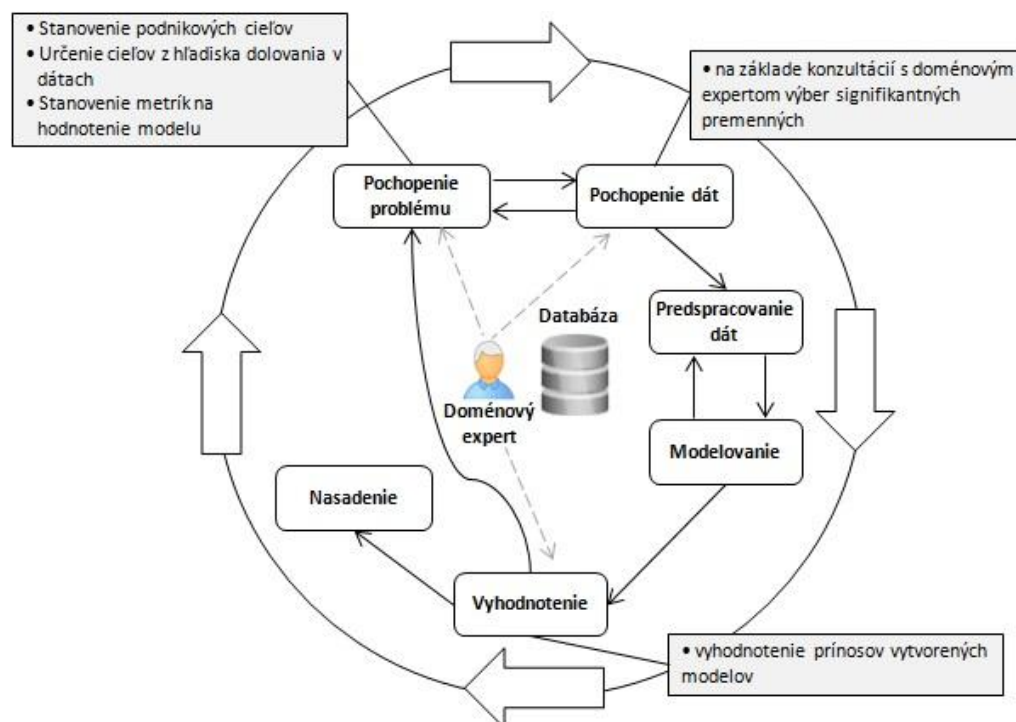
Prínos preukazuje i práca [10] z odvetvia obnoviteľných zdrojov energie, konkrétne veternej energie, kde sú použité techniky dolovania v dátach na predikciu porúch uhlíkových komponentov generátora veternej turbíny, pretože včasné predvídanie možnej chyby či poruchy umožňuje vykonať včasnú údržbu či výmenu určitých komponentov. Spoločnosť mala k dispozícii dáta z 27 veterných turbín zaznamenávané v 10 minútových intervaloch. Záznamy obsahujú viac ako 100 parametrov, ktoré môžeme zhruba rozdeliť na parametre vetra, výkonnostné parametre, parametre vibrácií a parametre teploty. Na modelovanie predikcie chyby v čase  $t+21$  boli použité 4 algoritmy – viacvrstvový perceptrón, boosting strom,  $k$ -NN (pre  $k = 10$ ) a metóda podporných vektorov, z ktorých najlepší vyšiel boosting strom (s počtom stromov 110).

#### D. Predikcia kvality

Príkladom pre predikciu je práca [2] z oblasti polovodičovej výroby. Dáta pozostávali zo 643 prípadov a 34 atribútov, pričom využili genetické programovanie (kontrolné parametre: maximálny počet generácií = 1000; veľkosť populácie = 50; jedno bodová úroveň kríženia = 0,7; úroveň mutácie = 0,044). Presnosť bola rátaná pomocou priemernej absolútnej percentuálnej chyby (dosiahnutá úroveň 95,89%).

### III. NÁVRH

Politikou výrobcu polyamidových vlákien, určených pre priemyselné použitie a ktorého dáta nám boli poskytnuté, je pre svojich odberateľov zaistenie štandardu vysokej kvality.



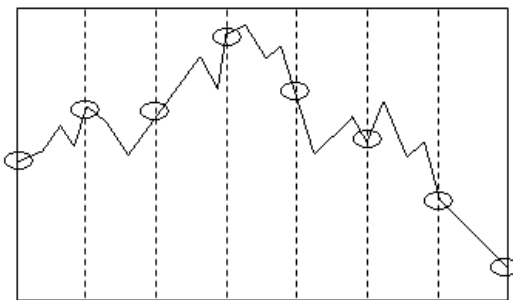
Obr. 1 CRISP-DM vo výrobnom prostredí

Cieľom tejto časti je preto z hľadiska dolovania v dátach navrhnúť ako definovať, čo ovplyvňuje a spôsobuje, že vyrobené cievky nespĺňajú požadovanú kvalitu. Za cieľový atribút môžeme zvoliť numerický atribút Počet chýb, ale aj symbolický atribút Trieda, pretože cievka je priradená do určitej triedy kvality na základe počtu chýb, ktoré sa na danej cievke vyskytli. Vzhľadom na to, môžeme danú úlohu riešiť tromi spôsobmi. Prvým je klasifikácia, kde je predikovaný atribút nominálny. V našom prípade by išlo o atribút Trieda, kde triedy 0 a 1 predstavujú kvalitnú cievku a triedy 2 a 3 nekvalitnú cievku. Druhú možnosť predstavuje predikcia. V tomto prípade by sme za cieľový atribút zvolili numerický atribút Počet chýb. Popisné dolovanie v dátach, spracované v prácach [18], [19] a [20] spolu s využitím asociačných pravidiel predstavuje tretiu možnosť.

Vstupný súbor obsahuje veličiny namerané zo senzorov počas výrobného procesu, ide o numerické atribúty. Hodnoty týchto atribútov sa zaznamenávajú každú minútu, a to 24 hodín denne. Dátum a čas zaznamenania týchto hodnôt je taktiež súčasťou vstupného súboru. Výstupný súbor obsahuje záznamy predstavujúce jednu vyrobenú cievku. Tento súbor tiež obsahuje numerické atribúty, ktoré vypovedajú, ako už bolo spomenuté vyššie, o kvalite vyrobenej cievky.

#### A. *Predspracovanie dát*

Ako je prezentované v [6], v prípade, že sú vybrané dáta príliš veľké, môžeme vybrať reprezentatívnu vzorku dát, a to napríklad pomocou metódy vzorkovania (viď Obr.2). Práve dáta z výrobných procesov sú charakteristické svojou veľkosťou, rozmernosťou a sú zoradené chronologicky [21]. Príkladom takýchto dát, sú dáta, ktoré obsahujú určité parametre výrobných zariadení (napr. teplota, tlak, a pod.) alebo namerané hodnoty vyrobeného zariadenia či výrobku.



Obr. 2 Príklad redukcie počtu hodnôt procesného parametra

Vo fáze predspracovania dát z výrobného procesu sa často stretávame s chýbajúcimi a prázdnyimi hodnotami, ktoré vznikajú pri prerušení výroby, alebo poruche senzorov z rozličných dôvodov. Pokiaľ sa chýbajúce hodnoty vyskytujú len v jednom alebo zopár záznamoch, je dobré tieto hodnoty nahradiť napríklad priemerom resp. modusom. V opačnom prípade sa odporúča dané hodnoty odstrániť, pretože ide o dlhší časový horizont počas ktorého sa nevyrábalo a nie len o krátkodobý výpadok meracieho zariadenia.

#### B. *Modelovanie*

V tejto fáze sa máme záujem sústrediť na využitie metód tvorby rozhodovacích stromov, ktorých algoritmy sú často vhodné tak na klasifikáciu, ako aj predikciu (napr. CART [12]). Taktiež je pre nás dôležité nielen získať model, ale aj vedieť získať pravidlá z neho v tvare zrozumiteľnom napr. pre operátora výroby.

#### C. *Vyhodnotenie a interpretácia*

Spôsobov vyhodnotenia je niekoľko. V prípade klasifikácie pôjde o presnosť, citlivosť a návratnosť. Pri prediktívnom dolovaní v dátach navrhujeme použiť ako metriku často používanú charakteristiku RMSE (z angl. Root Mean Squared Error), ktorá je daná vzťahom:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (1)$$

kde  $y_i$  je empirická hodnota a  $\hat{y}_i$  je teoretická hodnota pre  $i$ -ty príklad zo súboru  $n$  príkladov tvoriacich tréningové dáta [22]. Výsledné znalosti budeme prezentovať vo forme pravidiel, ktoré sú jednoduché pre pochopenie.

#### IV. ZÁVER

Idea vyhľadávania vzorov vo výrobe, podnikaní alebo medicínskych dát nie je nová. Uvedené práce poskytujú sumárny prehľad problematiky týkajúcej sa využitia metód dolovania v dátach vo výrobnom priemysle, ktoré vedú k zvýšeniu kvality finálnych produktov. Zo spomínaných úloh sme spracovali popis produktu za účelom zistenia dôležitých premenných, klasifikáciu a predikciu kvality. Ako je ale spomenuté aj v [23], výskumy v tejto oblasti generujú často pravidlá, kde sa neberie do úvahy ich kvalita a získane vzťahy sú príliš zložité na pochopenie. Naším zámerom je preto v budúcnosti zamerať sa na nájdenie vhodnej reprezentácie získaných znalostí, kde plánujeme využiť a porovnať rôzne metódy na indukciu rozhodovacích stromov.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

#### REFERENCIE

- [1] C. Çiflikli and E. Kahya-Özyirmidokuz, "Implementing a data mining solution for enhancing carpet manufacturing productivity," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 783–788, Dec. 2010.
- [2] T.-S. Li, C.-L. Huang, and Z.-Y. Wu, "Data Mining using Genetic Programming for Construction of a Semiconductor Manufacturing Yield Rate Prediction System," *J Intell Manuf*, vol. 17, no. 3, pp. 355–361, Jun. 2006.
- [3] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, R. Wirth, T. Reinartz, and C. Shearer, "CRISP-DM 1.0 - Step-by-step data mining guide." Aug-2000.
- [4] J. Paralič, *Objavovanie znalostí v databázach*. Košice: Elfa, 2003.
- [5] K. Wang, S. Tong, B. Eynard, L. Roucoules, and N. Matta, "Review on Application of Data Mining in Product Design and Manufacturing," in *Fourth International Conference on Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007*, 2007, vol. 4, pp. 613–618.
- [6] G. Köksal, İ. Batmaz, and M. C. Testik, "A review of data mining applications for quality improvement in manufacturing industry," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13448–13467, Sep. 2011.
- [7] N. Kitcharoen, S. Kamolsantisuk, R. Angsomboon, and T. Achalakul, "RapidMiner framework for manufacturing data analysis on the cloud," in *2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2013, pp. 149–154.
- [8] F. J. Martínez-de-Pisón, A. Sanz, E. Martínez-de-Pisón, E. Jiménez, and D. Conti, "Mining association rules from time series to explain failures in a hot-dip galvanizing steel line," *Computers & Industrial Engineering*, vol. 63, no. 1, pp. 22–36, Aug. 2012.
- [9] I. Tomek, "Two Modifications of CNN," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-6, no. 11, pp. 769–772, Nov. 1976.
- [10] A. Verma and A. Kusiak, "Fault Monitoring of Wind Turbine Generator Brushes: A Data-Mining Approach," *J. Sol. Energy Eng.*, vol. 134, no. 2, pp. 021001–021001, Feb. 2012.
- [11] M. Shabbaz, M. Srinivas, J. A. Harding, and M. Turner, "Product Design and Manufacturing Process Improvement Using Association Rules," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 220, no. 2, pp. 243–254, Jun. 2006.
- [12] D. T. Pham and A. A. Afify, "Machine-learning techniques and their applications in manufacturing," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 219, no. 5, pp. 395–412, May 2005.
- [13] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Taylor & Francis, 1984.
- [14] J. R. Quinlan, "Induction of Decision Trees," *MACH. LEARN*, vol. 1, pp. 81–106, 1986.
- [15] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.
- [16] Rulequest Research, "C5.0: An Informal Tutorial." 2013.
- [17] L. Rokach and O. Maimon, "Data Mining for Improving the Quality of Manufacturing: A Feature Set Decomposition Approach," *J Intell Manuf*, vol. 17, no. 3, pp. 285–299, Jun. 2006.
- [18] A. Lukáčová and J. Paralič, "Finding Patterns in Industrial Process Data," in *Electrical Engineering and Informatics V Proceeding of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice*, Košice, 2013.
- [19] A. Lukáčová and J. Paralič, "The use of data mining to identify patterns in the industrial process data," in *WIKT 2013*, Herľany, 2013, pp. 65–69.
- [20] A. Lukáčová, "A review of data mining applications in manufacturing," in *SCYR 2012 : proceedings from conference*, Herľany, 2012, pp. 64–67.
- [21] T. Fu, "A review on time series data mining," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 164–181, Feb. 2011.
- [22] E. Ostertagová, *Aplikovaná štatistika*. Košice: Equilibria, 2013.
- [23] J. A. Harding, M. Shabbaz, Srinivas, and A. Kusiak, "Data Mining in Manufacturing: A Review," *J. Manuf. Sci. Eng.*, vol. 128, no. 4, pp. 969–976, Dec. 2005.

# Možnosti využívania nizkopotenciálneho tepla

<sup>1</sup>Ivan Juhas, <sup>2</sup>Ján Tkáč

<sup>1</sup>Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>2</sup>Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>ivan.juhas@student.tuke.sk, <sup>2</sup>jan.tkac@tuke.sk

**Abstrakt**—Na zabezpečovanie svojich energetických potrieb v súčasnosti ľudstvo využíva v prevažnej miere energiu, uloženú vo fosílnych palivách. Fosílna palivá však patria medzi neobnoviteľné zdroje energie a ich zásoby sú obmedzené. So zmenšovaním ich zásob vyvstáva otázka náhrady tohto energetického zdroja. Jednu z alternatív predstavuje nizkopotenciálne teplo. Tento energetický zdroj sa nachádza v našom prostredí vo veľkom množstve, nakoľko vzniká ako koncová forma všetkých premien energie počas výroby a využívania. Jedná sa o energiu, ktorá je kvantitatívne mimoriadne výdatná a v termodynamickú sústavu Zem spôsobuje postupne globálne otepľovanie. Nizkopotenciálne teplo je však možné využívať transformáciou na vyšší potenciál pomocou zariadení nazývaných tepelné čerpadlá, zabezpečujúcich premenu kvantity na kvalitu. V súčasnosti prekonáva prudký vývoj najmä z hľadiska zvyšovania efektivity a kvality regulácie.

**Kľúčové slová**—nizkopotenciálne teplo, obnoviteľné zdroje energie, tepelné čerpadlo, vykurovací faktor

## I. ÚVOD

Napriek tomu že zákon zachovania energie hovorí, že energiu nevieme vyrobiť, stratiť ani zmariť neustále premieňame energiu koncentrovanú najčastejšie v chemických väzbách na inú, pre nás použiteľnú formu energie, najčastejšie elektrickú energiu a tepelnú. Konečným produktom premien energie je však nizkopotenciálne teplo (NPT), pre ktoré taktiež platí zákon o zachovaní energie a dosahuje vysokú kvantitatívnu úroveň [1],[2].

Najmä v súčasnej dobe je otázka využívania nizkopotenciálneho tepla mimoriadne aktuálna. Ceny momentálne najpoužívanejších zdrojov energie, medzi ktorými výrazne prevažujú fosílna palivá neustále rastú a ich zásoby sa znižujú. Potreba využívania obnoviteľných zdrojov energie je teda čoraz výraznejšia [3],[4].

Využívanie nizkopotenciálneho tepla predstavuje jednu z možností, ktorou je možné nahradiť energiu získavanú z neobnoviteľných zdrojov. Táto alternatíva bola dlhé obdobie prakticky nepovšimnutá, vzhľadom na technologickú náročnosť tohto procesu. Využívaniu NPT a iných obnoviteľných zdrojov energie určite neprospevala ani lobby presadzujúca využívanie fosílnych palív. S vedeckým pokrokom prišli metódy, ktoré umožňujú efektívne a cenovo prijateľné využívanie nizkopotenciálnej energie [5],[6].

Nizkopotenciálne teplo má svoj pôvod v slnečnej energii a geotermálnej energii, jeho zdrojom je taktiež ľudská činnosť. Akumuluje v prostredí v prevažnej miere vo vode pôde a vzduchu. Nizkopotenciálne teplo má vysokú kvantitu a výrazne ovplyvňuje termodynamickú rovnováhu na planéte. Z ekologického hľadiska vytvára nežiaduce tepelné emisie prispievajúce ku globálnemu otepľovaniu planéty.

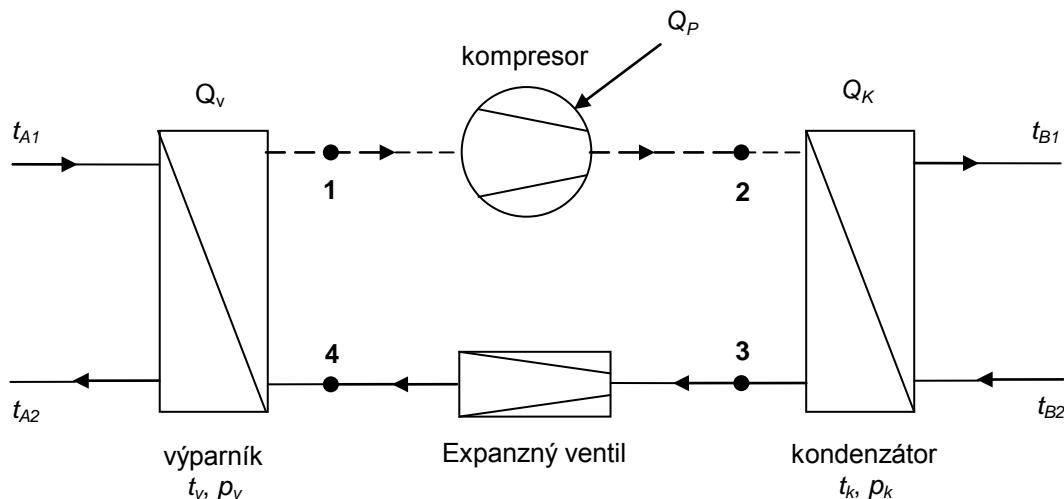
Vzhľadom na to, že nizkopotenciálne teplo nemá dostatočnú energiu na priame využitie, vyžaduje si aplikáciu sofistikovaných technológií, ktoré ho umožňujú transformovať na vyšší teplotný potenciál.

## II. TEPELNÉ ČERPADLO

Tepelné čerpadlo je tepelný stroj vhodný na využívanie nízkoenergetického tepla. Je to zariadenie, ktoré prečerpáva teplo z jedného miesta na druhé a zároveň ho prevádza na vyšší teplotný potenciál [7],[8].

Tepelné čerpadlá sa vyhotovujú najčastejšie ako kompresorové, alebo hybridné. Hybridné čerpadlá sú však menej rozšírené a vyhotovujú sa len na zákazku, čo ich predražuje. Sú vhodné pre veľké tepelné toky.

V súčasnosti sa najviac používajú kompresorové tepelné čerpadlá. Princíp kompresorového tepelného čerpadla je uvedený na Obr. 1.



Obr.1 Bloková schéma kompresorového tepelného čerpadla

Základom kompresorového tepelného čerpadla je chladivo, ktoré prechádza primárnym aj sekundárnym okruhom a výrazne ovplyvňuje efektivitu tepelného čerpadla. Chladivo expanduje v expanznom ventile. Dochádza k prudkému zníženiu tlaku a zároveň aj teploty. Chladivo ďalej prechádza do výparníka, kde odoberá teplo médiu privedenému do výparníka. Dochádza k odpareniu chladiva, ktoré je následne z výparníka odsaté kompresorom. V kompresore dochádza ku kompresii pár chladiva [8],[9]. Teplota chladiva sa prudko zvyšuje a dochádza k transformácii na vyšší teplotný potenciál. Z kompresora je chladivo vytlačené do kondenzátora, kde dochádza k odovzávaniu tepelnej energie chladiva médiu privedenému do kondenzátora. Zároveň dochádza pri kondenzačnom tlaku ku kondenzácii chladiva. Kvapalné chladivo sa zhromažďuje pred expanzným ventilom a expanduje do výparníka. Cyklus sa tým uzatvára.

Energetická náročnosť prečerpávania tepla sa vo všeobecnosti dá vyjadriť **vykurovacím faktorom**  $\varepsilon_T$ , inak nazývaným aj COP (coefficient of performance):

$$\varepsilon_T = \frac{Q_B}{A} \quad (-; J; J) \quad (1)$$

kde

$Q_B$  je teplo odovzdané médiu B

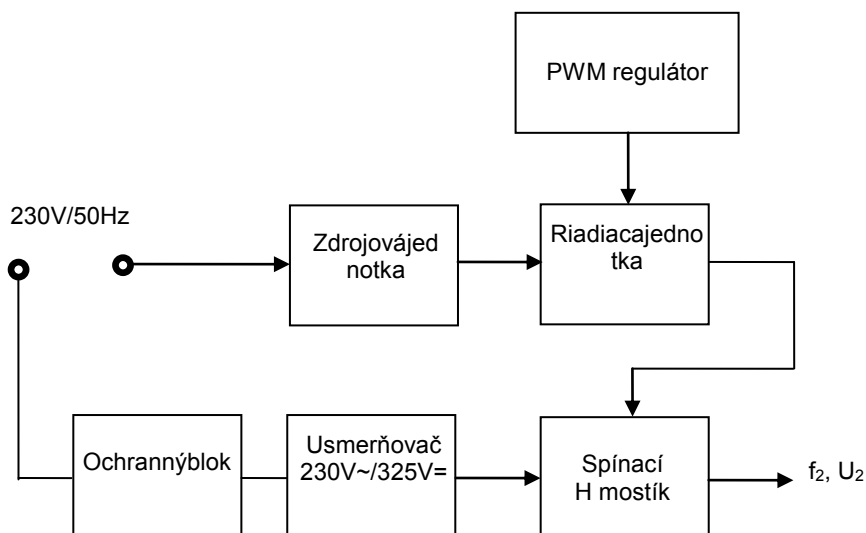
A dodaná pomocná energia

### III. EXPERIMENTÁLNA ČASŤ

V súčasnosti najčastejšie používané kompresorové tepelné čerpadlá majú pohon kompresora, ktorý nie je regulovaný a teda pracuje na nominálnych parametroch, ktoré však nemusia odpovedať aktuálnej potrebe dodávaného tepla. Takýto stav má nepriaznivý vplyv na účinnosť tepelného čerpadla. Okamžitá potreba produkovaného tepla tepelným čerpadlom sa neustále mení podľa dennej doby, ročného obdobia a tiež podľa požiadaviek spotrebiteľa.

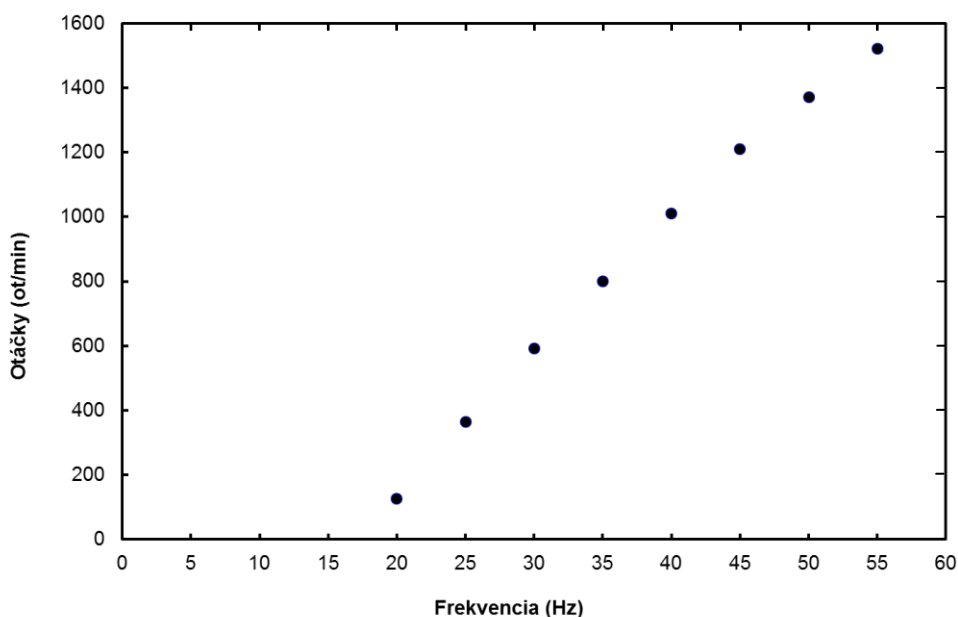
Za účelom experimentálneho overenia možnosti prispôsobenia výkonu tepelného čerpadla veľkosti okamžitej spotreby tepla bola navrhnutá riadiaca jednotka umožňujúca dosiahnutie stanoveného cieľa pomocou frekvenčnej regulácie pohonu kompresora. Táto regulácia mení výšku a frekvenciu napájacieho napätia v dôsledku čoho sa mení aj výkon kompresora.

Táto riadiaca jednotka je skonštruovaná podľa blokovej schémy na obr. 2.



Obr.2 Bloková schéma zdroja s premenlivou frekvenciou a napätím

Činnosť zariadenia bola najprv overená na pohone 1f asynchrónneho motora. Výsledky z overovania činnosti sú na Obr. 3.



Obr.3 Závislosť otáčok 1f asynchrónneho motora od frekvencie zdroja

Elektronické obvody zdroja sú realizované doske plošných spojov, ktorá je umiestnená v skrinke podľa Obr. 4. Na čelnom paneli je umiestnený vypínač a ručný regulátor výstupných parametrov. Zadný panel je na Obr. 5. a obsahuje sieťový prívod 230V/50Hz, ochranné tavné poistky a regulovanú výstupnú trojzásuvku.



Obr.4 Čelný panel zdroja s frekvenčnou reguláciou



Obr.5 Zadný panel zdroja s frekvenčnou reguláciou

#### IV. ZÁVER

Využívanie nízkopotenciálneho tepla je v súčasnej dobe aj napriek kvantitatívnym a kvalitatívnym ukazovateľom podceňované a málo využívané. Perspektíva využívania do budúcnosti je však nielen vysoká, ale aj nevyhnutná. Z hľadiska termodynamickej stability planéty je súčasný spôsob využívania energie vhodný nakoľko neustále sa vychádza z ťažby zdrojov energie, ich využitiu na produkciu žiadanej formy energie, pričom výsledný produkt sa zanedbáva a považuje sa za nízkopotenciálnu odpadovú energiu. Nerešpektuje sa pri tom zákon o zachovaní energie. Jedným zo zariadení určených na využívanie nízkopotenciálneho tepla je tepelné čerpadlo. Jeho princíp je známy už od roku 1852, ale v minulosti sa využíval len v oblasti chladiarenskej techniky. V súčasnosti sa však už tepelné čerpadlá začínajú presadzovať vo zvýšenej miere ako alternatíva ku klasickým zdrojom energie. Na demonštračné účely bolo na KEE skonštruované tepelné čerpadlo vzduch – voda, ktoré využíva fixný kompresor pracujúci stále na nominálnych parametroch. To spôsobuje, že v niektorých situáciách má tepelné čerpadlo nízky vykurovací factor. Takáto konštrukcia sa síce bežne používa u väčšiny tepelných čerpadiel, avšak najoptimálnejšie parametre sa dosahujú vtedy, ak množstvo vyprodukovaného tepla odpovedá okamžitej spotrebe. Takéto prispôsobenie je však možné len použitím frekvenčnej regulácie pohonnej jednotky kompresora.

#### POĎAKOVANIE

Táto práca bola realizovaná s podporou projektu Agentúry na podporu výskumu a vývoja APVV-0185-10.

#### LITERATÚRA

- [1] B. Hermanovsky, "Ivan: Energie pro 21. století," skriptá ČVUT 1992
- [2] I. Iliáš a kol. 2006, "Možnosti využívania slnečnej energie," Bratislava Energetické centrum Bratislava, 2006
- [3] Uniba [online]. [cit. 2014-02-08]. Dostupné na internete <<http://www.daa.fmph.uniba.sk/index.php?utid=64&id=110>>.
- [4] T. Matuška, "Alternativní zdroje energie," učebné texty ČVUT 2009
- [5] Euroweb [online] Dostupné na internete <[http://www.euroweb.sk/szchkt/tepelne\\_cerpadla/HPumps.htm#\\_Toc34983630](http://www.euroweb.sk/szchkt/tepelne_cerpadla/HPumps.htm#_Toc34983630)>.
- [6] "Stratégia vyššieho využitia obnoviteľných zdrojov energie v SR," Materiál ministerstva hospodárstva SR 2007
- [7] R. Karlik, "Tepelné čerpadlo pro váš dům," Praha 2009. 112 s. ISBN 978-80-247-2720-2
- [8] Transformacni-technologie [online] Dostupné na internete <<http://www.transformacni-technologie.cz/>>.
- [9] V. Havel'sky, "Chladivá," Slovenský zväz pre chladiacu a klimatizačnú techniku



# Meranie vlastností solárnych systémov polohovaných smerom k Slnku

<sup>1</sup>Samuel KAZÁR, <sup>2</sup>Ján TKÁČ

<sup>1</sup>Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>2</sup>Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>samuel.kazar@student.tuke.sk, <sup>2</sup>jan.tkac@tuke.sk

**Abstrakt**—Za účelom dosiahnutia čo najväčšej efektivity pri použití solárnych systémov je vhodné realizovať prevádzkové merania umožňujúce sledovať a optimalizovať prevádzku. Článok sa zaoberá problematikou návrhu a realizácie nízkonákladového univerzálneho meracieho systému pre experimentálne merania. Na dosiahnutie tohto cieľa bola použitá elektronická platforma Arduino založená na flexibilnom a voľne dostupnom technickom a programovom vybavení.

**Kľúčové slová**—merací systém, experimentálne meranie, Arduino UNO, solárny panel

## I. ÚVOD

Za účelom dosiahnutia maximálneho energetického zisku slnečných kolektorov je potrebné brať do úvahy orientáciu a polohu kolektorov voči Slnku. Pri projektovaní solárnych systémov musíme vychádzať z poznatkov o dostupnosti slnečnej energie v danej lokalite. Údaje o množstve slnečného žiarenia dopadajúceho na povrch kolektorov môžeme získať uskutočnením meraní pyranometrami, alebo prostredníctvom analytických výpočtov. Pomocou takto získaných údajov vieme slnečné kolektory umiestniť na vhodné miesto, ale aj napriek tomu je pre dosiahnutie maximálneho energetického zisku nutné meniť ich polohu. Polohovacie zariadenia, ktoré slúžia na tento účel nazývame trackery. Slnečné kolektory vykazujú najvyššiu účinnosť a energetický zisk, keď je kolektorová rovina v pravom uhle voči dopadajúcemu slnečnému žiareniu. Polohovacie zariadenia sú pre solárne systémy slnečných elektrární s vysokým stupňom koncentrácie slnečného žiarenia nevyhnutnosťou. Pri rovinných slnečných kolektoroch a fotovoltaických článkoch, ktoré majú vysoký zberný uhol slnečného žiarenia nie je polohovanie nutnosťou. Polohovacie zariadenie menia svoju polohu na základe údajov, ktoré sú získavané zo solárnych, alebo svetelných senzorov, prípadne na základe GPS polohy, alebo pomocou výpočtových programov inštalovaných v riadiacich jednotkách. Riadiaca jednotka tieto údaje vyhodnocuje a ďalej dáva inštrukcie pre krokový, alebo servomotor, ktorý vykoná pohyb trackeru. Uhol sklonu (eleváciu) je potrebné nastaviť podľa zemepisnej šírky daného miesta. Tracker je to vlastne mechanická konštrukcia, ktorá sa pohybuje v priebehu dňa rovnomerne, alebo vo vopred stanovených časových intervaloch s pohybom slnka. Solárne moduly sú natáčané do presných pozícií pre maximalizovanie energetického zisku. V porovnaní s pevným nastavením zvyšujú polohovacie zariadenia energetický zisk približne o 30 %.[1]

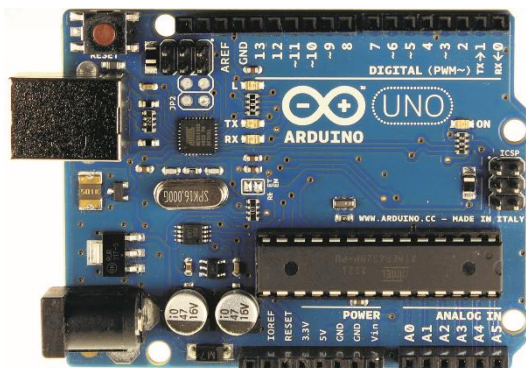
Pre realizáciu meracieho systému, ktorý bude slúžiť na meranie parametrov solárnych systémov existuje v dnešnej dobe viacero možností ako napríklad použitie programovateľných logických automatov, platformy CompactRio s vývojovým prostredím LabView, platformy RaspberryPi atď. Pre realizáciu nízkonákladového meracieho systému bola zvolená platforma Arduino. Realizovaný merací systém umožňuje meranie teploty, napätia a prúdu a prešiel testom funkčnosti pri experimentálnom meraní.

## II. HARDVÉROVÉ RIEŠENIE S POUŽITÍM PLATFORMY ARDUINO UNO

*Arduino UNO je voľne šíriteľný projekt založený na elektronickej platforme. Voľne dostupné je aj vývojové prostredie a zdrojové kódy knižnic pre prácu s Arduinom. Základom je vývojová doska (plošný spoj - hardvér) a vývojové prostredie (softvér). Arduino môže získavať údaje z vonkajšieho prostredia prostredníctvom rôznych senzorov, na základe ktorých je potom schopné ovládať iné zariadenia ako*

*napríklad motory, relé, osvetlenie, atď. Arduino projekt môže fungovať samostatne, alebo môže byť ovládaný aplikáciou z počítača, tabletu, alebo mobilného telefónu. Pri jednoduchých projektoch nie je nutné elektronický obvod spájať, ale iba pospájať na kontaktnom poli. Potrebujeme k tomu iba Arduino vývojovú dosku (*

Obr. 1), USB kábel na pripojenie k počítaču, kontaktné pole, elektronické súčiastky a nainštalované vývojové prostredie [2].



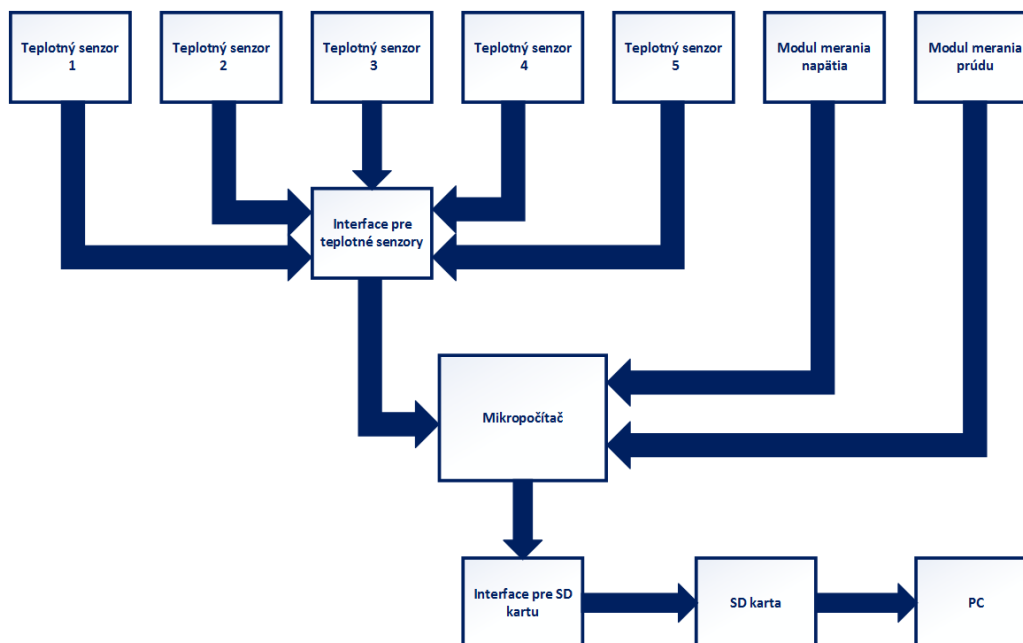
Obr. 1 – Vývojová platforma Arduino

Základom zariadenia je 8-bitový mikrokontrolér od firmy Atmel. Tento mikrokontrolér kombinuje 32 KB ISP flash pamäte, 1024B EEPROM, 2 KB SRAM, 23 vstupno-výstupných programovateľných liniek, 32 univerzálnych registrov, tri flexibilné časovače/čítače, synchronno/asynchronné sériové rozhranie, sériové periférne rozhranie, 6-kanálový 10-bitový A/D prevodník, programovateľného „strážcu“ s vnútorným oscilátorom a päť softvérovo voliteľných režimov úspory energie. Kontrolér pracuje v rozmedzí 1,8-5,5 V.

Programovací jazyk pre uvedené zariadenie vychádza z jazyka C/C++ ku ktorému sú k dispozícii knižnice programov. Vývojové prostredie Arduino je bezplatne šírený produkt, ktorý je možné používať bez akýchkoľvek obmedzení [3].

### III. VLASTNOSTI MERACIEHO SYSTÉMU

Bloková schéma meracieho systému je na Obr.2. Signály z teplotných senzorov sú prepojené do modulu rozhrania pre teplotné senzory, ktorý je pripojený k základnej doske zabezpečujúcej spracovanie prichádzajúcich signálov zo senzorov. Meranie teplôt je realizované prostredníctvom 1-Wire technológie. Modul merania napätia a prúdu je vzhľadom na počet senzorov k základnej doske pripojený priamo. Vyhodnotenú signály sú zaznamenávané na SD kartu, z ktorej je možné údaje ďalej spracovávať v exteriernom počítači.



Obr. 2 –Bloková schéma meracieho systému

#### A. Teplotný senzor

Použitý bol digitálny teplotný senzor, ktorý komunikuje cez zbernicu 1-Wire. Výhodou tejto technológie je, že na komunikáciu je potrebný iba jeden vodič. Pracuje s napätím 2,8-6 V a využíva CMOS/TTL logiku. Každé 1-Wire zariadenie má svoju jedinečnú 64-bitovú adresu, ktorú využíva na identifikáciu na zbernici, čo umožňuje pripojenie viacerých senzorov na jednu zbernicu. Skladá sa z 8 bitov, ktoré identifikujú typ zariadenia, 48 bitov pre sériové číslo a 8 bitov pre kontrolu cyklickým kódom CRC. Rozsah meraných teplôt je v rozsahu od -55 °C do +125 °C s presnosťou  $\pm 0,5$  °C v rozsahu od -10 °C do +85 °C.[4]

#### B. Modul merania napätia

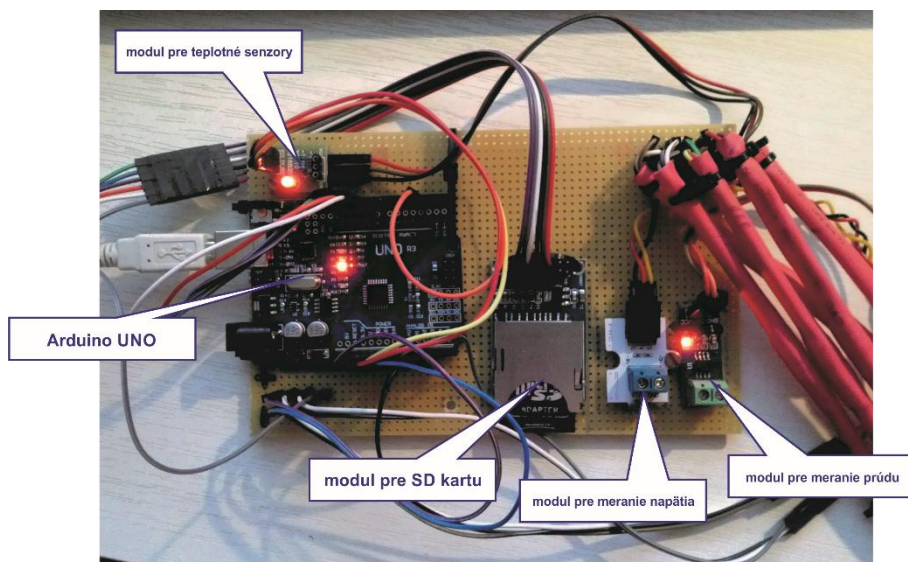
Senzor pre meranie napätia je analógový senzor, ktorý pracuje na princípe odporového deliča. Vstupné napätie dokáže znížiť 16-krát. Maximálna hodnota napätia na analógovom vstupe Arduina je 5 V, takže maximálne merané napätie môže byť  $16 \times 5 \text{ V} = 80 \text{ V}$ . Minimálna hodnota napätia, ktoré dokáže senzor zachytiť je 0,02445 V. Napájacie napätie senzora môže byť v rozsahu 3,3-5 V [5].

#### C. Modul merania prúdu

Prúdový senzor je určený na presné meranie jednosmerného, alebo striedavého prúdu. Je konštruovaný pre meranie prúdov v rozsahu  $\pm 30 \text{ A}$  s citlivosťou 66 mV/A. Napájacie napätie senzora je 5 V. Senzor je realizovaný na základe Hallovoho javu, ktorý je ovplyvňovaný vonkajším magnetickým poľom vznikajúcim v blízkosti vodiča, ktorým preteká prúd [6].

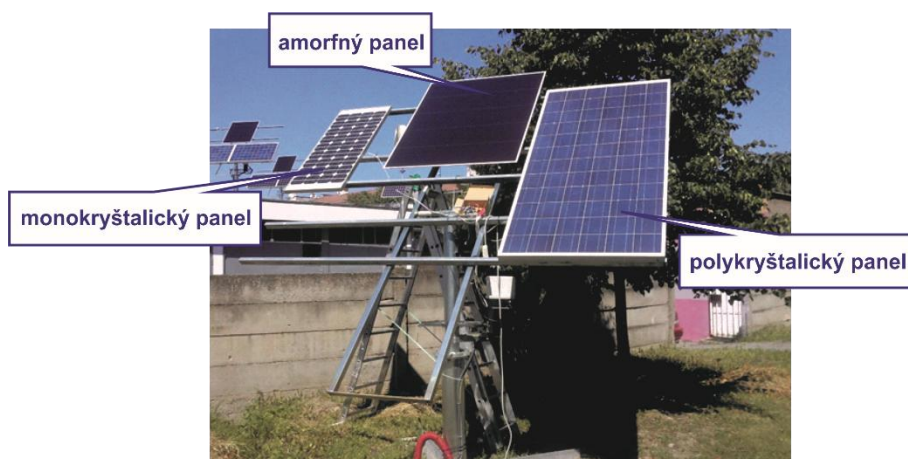
### IV. EXPERIMENTÁLNE MERANIE

Experimentálne meranie bolo zamerané na overenie činnosti teplotných senzorov pri meraní teploty solárnych panelov, ktoré so stúpajúcou teplotou vykazujú nárast sériového odporu a pokles účinnosti. Pri experimentálnom meraní boli realizovaným meracím systémom na Obr.3 merané teploty na povrchu solárnych fotovoltaických panelov. Teplotné senzory boli umiestnené v smere zvislej osi v 1/3 a 2/3 výšky.



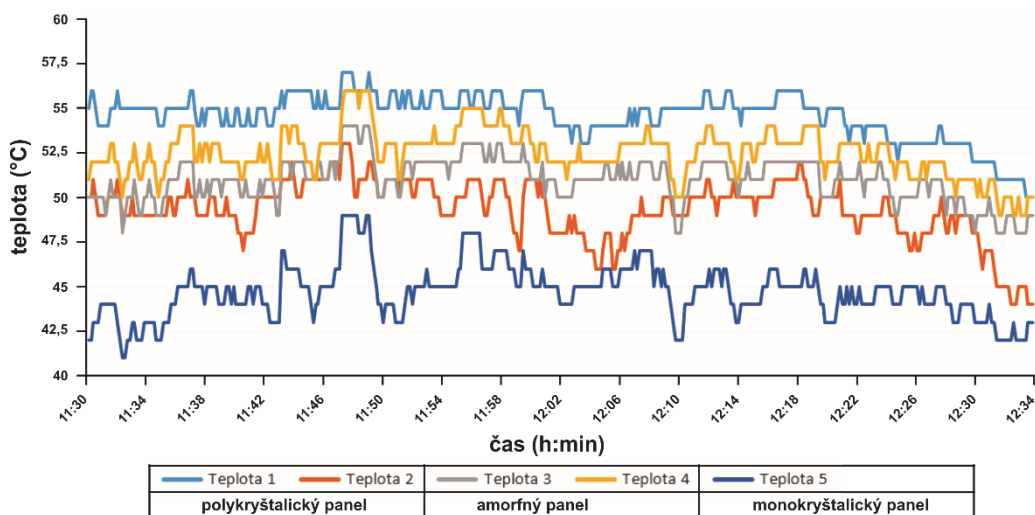
obr. 3 – realizovaný merací systém

Teploty boli merané na troch druhoch fotovoltaických panelov – monokryštalickom, polykryštalickom a amorfnom pre porovnanie hodnôt teploty v závislosti od druhu panelu Obr.4.



Obr. 4 – Merané solárne panely na polohovacom zariadení

Namerané hodnoty z experimentálneho merania sú zobrazené na Obr.5. Teplotné senzory boli umiestnené zo zadnej strany solárnych panelov. Z nameraných závislostí môžeme konštatovať, že pri rovnakých poveternostných podmienkach dosahoval najnižšie teploty monokryštalický panel - v priemere 44,69 °C. Naopak najvyššie teploty boli namerané na polykryštalickom panely - v priemere 51,95 °C. Amorfný panel dosahoval priemerné teploty 51,73 °C, čo je porovnateľné s polykryštalickým panelom. Meranie bolo uskutočnené 26.05.2014 od 11:30 hod v desaťsekundových intervaloch. Vykonaných bolo 392 cyklov merania.



Obr. 5 –Priebeh teploty na solárnych paneloch

## V. ZÁVER

Realizovaný nízkonákladový merací systém je určený na meranie prevádzkových parametrov solárnych systémov pre edukačné a vedeckovýskumné účely. Bol realizovaný s použitím elektronickej platformy Arduino, ktorá je založená na flexibilnom a voľne dostupnom technickom a programovom vybavení s prijateľnými finančnými nákladmi a plne postačovala pre realizovanie požadovaných meraní. Merací systém obsahuje senzory pre meranie teploty, modul pre meranie napätia a modul pre meranie prúdu. Nakoľko realizovaný merací systém pre meranie prevádzkových parametrov solárnych systémov je veľmi flexibilný a umožňuje pripojiť ďalšie senzory, bolo by do budúcnosti vhodné doplniť o dva meracie moduly napätia a dva meracie moduly prúdu, aby bolo možné realizovať merania na všetkých troch druhoch solárnych panelov naraz a tak vyhodnocovať merané údaje, kontinuálne pri meniacich sa meteorologických podmienkach a intenzite slnečného žiarenia.

## POUŽITÁ LITERATÚRA

- [1] J. TKÁČ –M. HVIZDOŠ: Polohovanie solárnych zariadení. In: Elektroenergetika, Vol.4, No.1 2011, s. 23-25. Dostupné na internete: <<http://jeen.fei.tuke.sk/index.php/jeen/article/view/203/177>> ISSN 1337-6756
- [2] M. BOŽÍK: Ako naprogramovať Arduino bez predchádzajúcich znalostí. [online]. [cit. 2014-07-12]. 2013. Dostupné na internete: <<http://mirobozik.sk/>>
- [3] Atmel 8-bit Microcontroller with 4/8/16/32KBytes In-System Programmable Flash. [online]. [cit. 2014-07-12]. 2013. Dostupné na internete: <[http://www.atmel.com/images/atmel-8271-8-bit-avr-microcontroller-atmega48a-48pa-88a-88pa-168a-168pa-328-328p\\_datasheet\\_summary.pdf](http://www.atmel.com/images/atmel-8271-8-bit-avr-microcontroller-atmega48a-48pa-88a-88pa-168a-168pa-328-328p_datasheet_summary.pdf)>
- [4] DS18B20 Programmable Resolution 1-Wire Digital Thermometer. [online]. [cit. 2014-07-12]. 2008. Dostupné na internete: <<http://datasheets.maximintegrated.com/en/ds/DS18B20.pdf>>
- [5] Octopus Analog Voltage Divider Brick. [online]. [cit. 2014-07-12]. 2013. Dostupné na internete: <[http://www.electronics.com/wiki/index.php?title=Octopus\\_Analog\\_Voltage\\_Divider\\_Brick](http://www.electronics.com/wiki/index.php?title=Octopus_Analog_Voltage_Divider_Brick)>
- [6] ACS712 Fully Integrated, Hall Effect-Based Linear Current Sensor IC with 2.1 kV RMS Isolation and a Low-Resistance Current Conductor. [online]. [cit. 2014-07-12]. 2006-2013. Dostupné na internete: <<http://www.allegromicro.com/~media/Files/Datasheets/ACS712-Datasheet.ashx>>



# Výbojové javy v atmosfére

<sup>1</sup>Jaroslav Vilim, <sup>2</sup>Ján Tkáč

<sup>1</sup> Department of Electric of Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>2</sup> Department of Electric of Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>jaroslav.vilim@student.tuke.sk, <sup>2</sup>jan.tkac@tuke.sk

**Abstrakt** — Výbojová činnosť v atmosfére svojimi výraznými optickými akustickými a deštruktívnymi efektmi, sústreďuje pozornosť obyvateľstva našej planéty od nepamäti. Vzniká najčastejšie v súvislosti s búrkovou činnosťou. Táto každoročne narastá a s ňou sa zvyšuje aj počet bleskov ktoré zasiahnu zem. Je zaznamenávaný aj nárast škôd na životoch a majetku. Vzhľadom na tento fakt vznikajú nové druhy zachytávacích zariadení, buď klasický systém zachytávačov ktorým sa mení tvar alebo aktívne zachytávače, u ktorých sa mení tvar aj spôsob vytvárania včasného protismerného výboja. V tomto článku sa bude pojednávať o zachytávacej elektróde metličkovitého tvaru.

**Kľúčové slová** — aktívny bleskozvod, blesk, hrom, ochrana, zachytávacie zariadenie

## I. ÚVOD

Výbojové procesy v atmosfére Zeme boli predmetom záujmu nie len v dôsledku výrazných optických a akustických efektov, ale aj významných deštruktívnych schopností a ohrozenia obyvateľstva. Atmosférická elektrina poskytovala aj možnosť vytvorenia atmosférických zdrojov vysokého napätia, ktoré sa používali v prvých štádiách výskumu. Výskyt výbojovej činnosti je však výrazne ovplyvnený procesmi v atmosfére Zeme, ktoré majú rozhodujúci vplyv na biosféru planéty.

Vznikom a výskytom atmosférickej elektriny sa zaoberajú vedci už celé desaťročia a aj napriek tomu neboli odkryté všetky princípy a javy, ktoré sprevádzajú atmosférické výboje. Najpreskúmanejší z atmosférických výbojov je blesk typu mrak – zem, ktorý spôsobuje najväčšie škody na životoch a na majetku.

Počet bleskov na celej planéte ktoré zasiahnu povrch Zeme, dosahuje cca 8 640 000 za deň. Napriek tomu, že je táto početnosť pomerne vysoká, je zaznamenaná neustále stúpajúca tendencia výskytu bleskov. Preto je výskum atmosférických výbojov, vývoj zachytávacích zariadení a metód ochrany vo všetkých odvetviach mimoriadne aktuálny.

Vhodnosť jednotlivých zachytávacích zariadení sa overuje v špeciálne zariadených laboratóriách, ktoré simulujú atmosférické výboje. V laboratórnych podmienkach však nie je možné dosiahnuť tak širokú škálu namáhania skúmaných zariadení ako je tomu v praxi. Z analýzy súčasného stavu vyplýva, že mechanizmus vzniku atmosférických výbojov a vhodné spôsoby ochrany pred atmosférickými výbojmi, ako aj overenie vhodnosti používania rôznych druhov zberacích zariadení je stále aktuálny a výskum sa orientuje na optimalizáciu tvaru zberacích elektród pre špeciálne podmienky a aplikácie. Zaujímavá je a j orientácia na zberacie systémy s viacerými hrotovými elektródami usporiadanými v tvare metličky.

## II. FORMY ATMOSFÉRICKEJ ELEKTRINY

Najvýraznejším prejavom výskytu atmosférickej elektriny je vznik výboja vo forme blesku. *Blesk je viditeľný svetelný jav, ktorý je sprevádzaný náhlym elektrickým výbojom prejavujúcim sa intenzívnym vyžarovaním svetla z výbojového kanála ako aj generovaním akustických vln nazývaných aj ako hrom.*

Veľmi závažné sú následky v mieste zásahu objektov na povrchu Zeme, často sprevádzaných vznikom požiaru. Atmosférická elektrina však nespôsobuje vznik výboja zasahujúci povrch

zeme ale prejavuje sa rôznymi optickými a elektromagnetickými efektmi. Najčastejšie sa tieto efekty prejavujú počas búrkovej činnosti. Elektrický výboj (blesk) vzniká len medzi centrami elektrického náboja opačnej polarity.

Poznáme nasledovné druhy výbojov, medzi kladným a záporným centrom:

- vo vnútri oblaku (IC)
- medzi blízkymi oblakmi (medzioblačný výboj) (CC)
- medzi oblakom a voľnou atmosférou (CA)
- medzi oblakom a zemou (CG) – (Obr. 1)
- medzi zemou a oblakom (GC) – (Obr. 2)



Obr. 1 Páskový a perlový blesk - mračno zem.



Obr. 2 Blesk rozvíjajúci sa zo zeme k mračnu.

Základné parametre záporného blesku druhu mrak zem (CG) sú v (tabuľka 1).

Tabuľka 1  
Parametre blesku

Parameter	hodnota
Potenciálový rozdiel	100 MV
Celkový prenesený prúd	30 kA
Teplota blesku	30 000 °C
Teplota na povrchu slnka	6 000 °C
Rýchlosť blesku	10 000 – 100 000 km/s
Počet výbojov v jednom priebehu	3 – 5
Priemerný interval medzi výbojmi	60 ms
Celková doba trvania	200 – 300 ms
Celkový prenesený náboj	20 – 100 C
Celková prenesená energia	$10^9 - 10^{10}$ J

Atmosférické výboje sa prejavujú aj vo forme ktoré sú doprevádzané iba oprickými efektmi a sú to polárna žiara – (Obr. 3), eliašov oheň, red sprites, blue jets, elves.



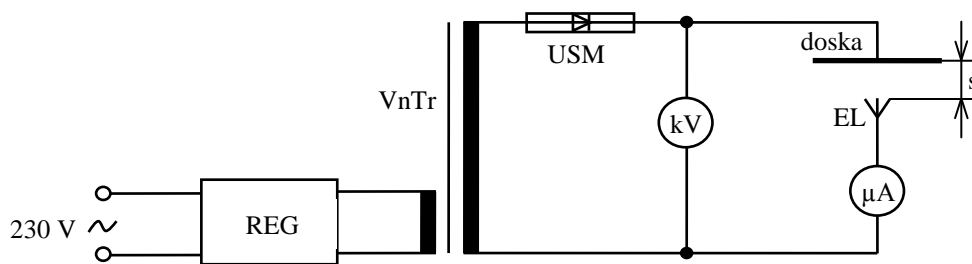
Obr. 3 Polárna žiara.

### III. EXPERIMENTÁLNE MERANIE

Experimentálne meranie bolo uskutočnené podľa schémy zapojenia na (Obr. 4). Cieľom bolo realizovať merania na elektródovom systéme doska - zachytávacie zariadenie. Zachytávacie zariadenie bolo realizované s elektródou s viacerými hrotmi. Pri meraní sa použili zachytávacie zariadenia s rôznou konfiguráciou elektród podľa (Obr 6).

Pri správnej činnosti bleskozvodov je vznik koróny rozhodujúcim aspektom preto sa pri meraniach sledovala veľkosť korónového napätia a prúdu. Koróna vzniká pri silne zakrivených elektródach a predstavuje neúplný samostatný výboj, ktorý sa objavuje pri počiatočnom napätí, ktoré je omnoho nižšie ako elektrická pevnosť elektródového usporiadania. Koróna je sprevádzaná zvukovým ako aj svetelným efektom (sršanie, svetielkovanie na špičke hrotu). Čím sa skôr začne ionizovať vzduch okolo hrotu zberača prostredníctvom koróny, tým sa skôr vytvorí vzostupný výboj smerom k blesku. Vďaka čomu sa zvyšuje šanca zachytenia blížiaceho sa blesku a jeho bezpečné zvedenie do zeme. Merania prebehali pri kladnom aj zápornom napätí na zachytávacom zariadení.



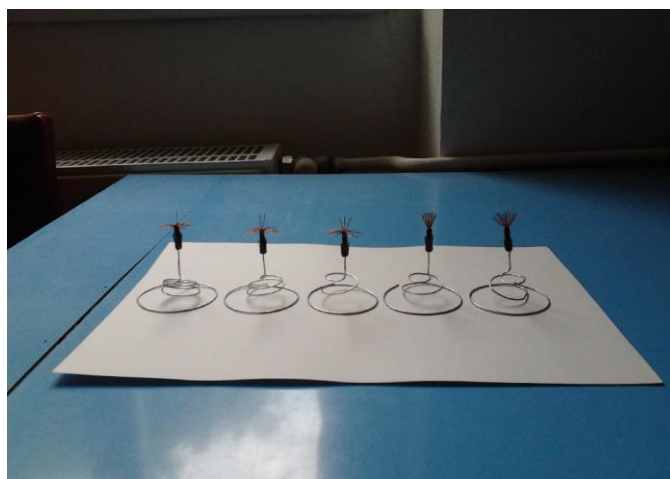


Obr. 4 Schéma zapojenia.

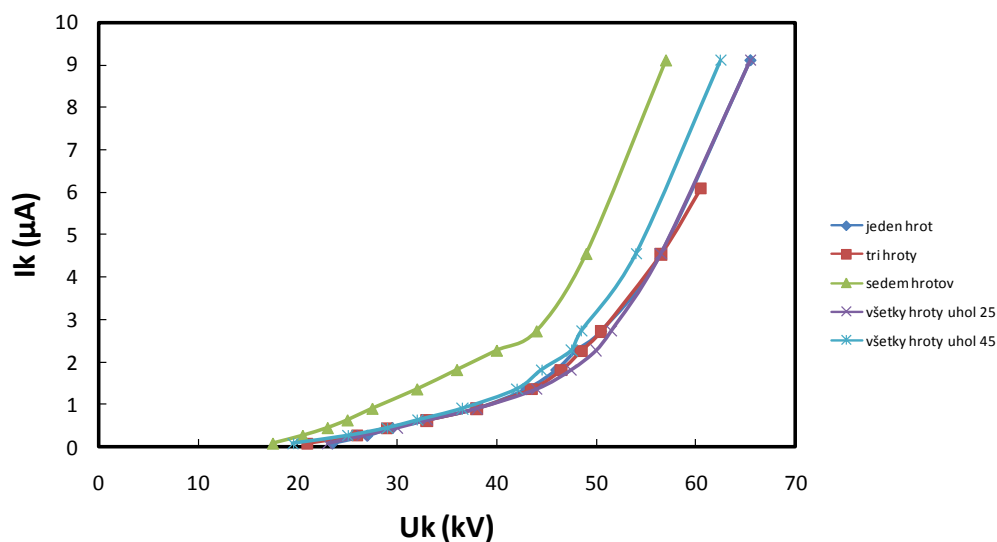
REG – regulátor  
VnTr – vysokonapäťový transformátor  
USM – vysokonapäťový usmerňovač  
kV – kilovoltmeter  
EL – hrotová elektróda  
s – preskoková vzdialenosť  
 $\mu\text{A}$  - miliampérmeter



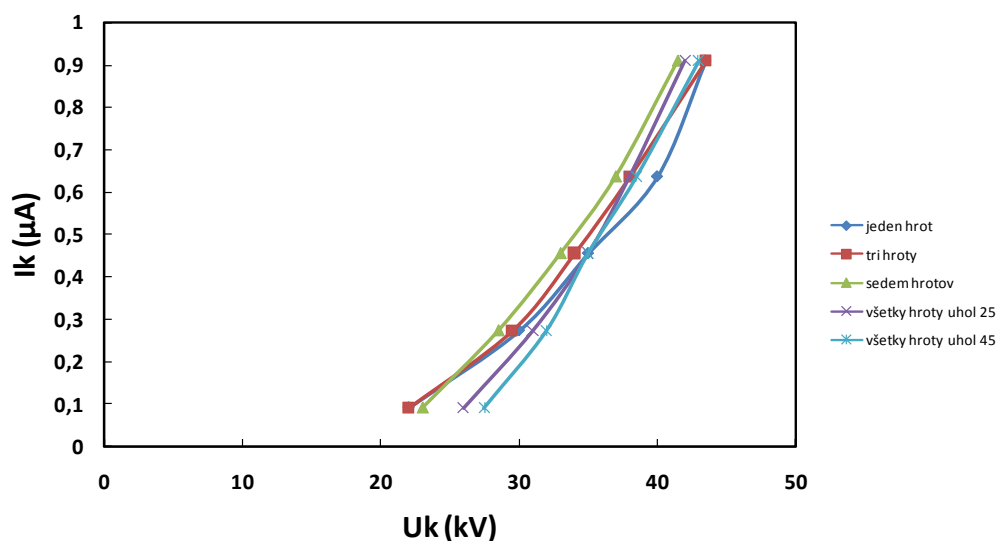
Obr. 5 Elektródový systém simulujúci usporiadanie zachytávacie zariadenie a mračno.



Obr. 6 Druhy použitých elektród.



Obr. 7 V-A charakteristiky korónových prúdov záporných elektród.



Obr. 8 V-A charakteristiky korónových prúdov kladných elektród.

Z charakteristík na (Obr. 7) vyplýva, že korónový prúd postupne rastie s rastúcim počtom hrotov. Korónový prúd na elektróde s tromi hrotmi stúpol o 2,56% oproti elektróde s jedným hrotom pri napätí 55kV a na elektróde so siedmimi hrotmi bol korónový prúd väčší až o 105,13%. Elektróda s hrotmi usporiadanými do tvaru kužeľa s vrcholovým uhlom 25° mala pri napätí 55 kV rovnaký korónový prúd ako elektróda s jedným hrotom a na elektróde s hrotmi usporiadanými do tvaru kužeľa s vrcholovým uhlom 45° bol korónový prúd väčší o 28,21% oproti elektróde s jedným hrotom. Namerané závislosti boli exponenciálneho charakteru.

Z charakteristík na (Obr. 8) vyplýva, že korónový prúd taktiež postupne rastie s rastúcim počtom hrotov ale nie tak výrazne. Na elektróde s tromi hrotmi stúpol korónový prúd o 15,63% oproti elektróde s jedným hrotom pri napätí 40 kV a na elektróde so siedmimi hrotmi stúpol korónový prúd o 28,13%. Na elektróde s hrotmi usporiadanými do tvaru kužeľa s vrcholovým uhlom 25° bol korónový prúd menší oproti hrotu so siedmimi hrotmi, ale oproti elektróde s jedným hrotom bol korónový prúd väčší o 20,31%. Na elektróde s hrotmi usporiadanými do tvaru kužeľa s vrcholovým uhlom 45° opäť poklesol korónový prúd, avšak oproti elektróde s jedným hrotom bol väčší o 14,06% pri napätí 40 kV.

#### IV. ZÁVER

Skúmanie vznik a pôvod atmosférickej elektriny a výskytu výbojovej činnosti v atmosfére

v podobe blesku a iných fyzikálnych prejavov, je v dôsledku dosahovaných účinkov veľmi dôležité. Deje prebiehajúce v atmosfére sú veľmi zložité. Znalosti jednotlivých fáz rozvoja blesku sú podstatné pre vývoj a konštrukciu zberacích zariadení.

V súčasnosti sa do popredia dostávajú aktívne bleskozvody ktoré boli zavedené v niektorých štátoch do bežného používania, čo bolo dosiahnuté podporou v rámci normotvornej činnosti. V rámci experimentálnych meraní ktoré boli realizované v spolupráci s výrobcami, bolo preukázané, že pri ich správnej činnosti je možné dosiahnuť významné finančné úspory, v dôsledku zníženia nákladov na bleskozvodovú sústavu hlavne na historických objektoch. Avšak v odbornej komunite je často účinnosť aktívnych bleskozvodov spochybňovaná a to aj napriek tomu, že vo väčšine krajín boli prijaté normy na ich používanie.

Uskutočnené merania poukázali na skutočnosť, že tvar elektród a im odpovedajúca nehomogenita poľa ovplyvňuje elektrickú pevnosť, pričom sa uplatňujú aj ultrakorónové efekty. Na základe zistených skutočností je možné konštatovať, že perspektívne bude potrebné zvýšiť presnosť konštrukcie a tvarový sortiment elektród s definovanou nehomogenitou. Experimentálne merania korónového prúdu boli realizované za účelom získania základných skúseností pre prípravu ďalších experimentov v tejto oblasti.

Jedným z najdôležitejších aspektov je zvýšiť ochranu zariadení a objektov pred atmosférickými výbojmi. Pri vývoji bleskozvodov bolo realizovaných a skúmaných množstvo konštrukcií elektródových systémov, pričom sa vyskytoval aj bleskozvod zakončený súborom hrotov – metličkový bleskozvod. V tejto práci sme sa zamerali na overenie vhodnosti používania takéhoto bleskozvodu, ktorý bol v konečnom dôsledkom nahradený tyčovým zberacím zariadením. Naše merania nepotvrdili vhodnosť použitia viacerých hrotových elektród na špičke zberacieho zariadenia, ani z hľadiska preskokového napätia, ani z hľadiska jednoduchosti konštrukcie. Ich výhodnosť je možné predpokladať v prípade aktívnych bleskozvodov vyžadujúcich masívnu ionizáciu prostredia pre vznik ústretového výboja.

#### POĎAKOVANIE

Táto práca vznikla za podpory vedeckej grantovej agentúry ministerstva školstva SR a Slovenskej akadémie vied v rámci projektu VEGA 1/0487/12 a APVV-0185-10.

#### LITERAÚRA

- [1] K. Marton, Technika vysokých napätí I. Bratislava: ALFA, 1984. 206s. ISBN 63-764-83,
- [2] J. Dudáš, Poznámky k aktívnym hromosvodům. In: ELEKTRO, roč. 7, 2006, č. 05, s. 45 – 9,
- [3] STN EN 62305-1: 2007, Ochrana pred bleskom. Časť 1: Všeobecné princípy,
- [4] J. Burant, Blesk a přepětí. Ústí nad labem: Horák, 2006. 256s. ISBN 80-86534-10-3,
- [5] STN EN 62305-3: 2007, Ochrana pred bleskom. Časť 3: Ochrana stavieb a ohrozenie života,
- [6] P. Šandrik, Technika vysokých napätí. Bratislava: STU, 2004. 178s. ISBN 80-227-2137-9,
- [7] R. Kvak, Elektrické javy v atmosfére Zeme. In: Slnko a vzduch: Zborník prednášok zo seminára pre učiteľov základných a stredných škôl, Bratislava: SAV, 2012. s. 33 – 37. ISBN 978-80-85754-25-4,
- [8] J. Bednář, Pozoruhodné jevy v atmosféře. Praha: ACADEMIA, 1989. 240 s. ISBN 80-200-0054-2,
- [9] A. Veverka, Technika vysokých napětí. Praha: ALFA, 1978. 296 s. ISBN 04-512-78.

# Digitálna Vodotlač v Transformačnej Oblasti vo Videu

<sup>1</sup>Vladimír Bánoci, <sup>1</sup>Gabriel Bugár, <sup>1</sup>Dušan Levický

<sup>1</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>vladimir.banoci@tuke.sk, gabriel.bugar@tuke.sk, dusan.levicky@tuke.sk

**Abstract** — Technika digitálnych vodoznakov predstavuje účinný prostriedok na ochranu informačného obsahu multimediálnych dát. Následkom toho informačný obsah označených multimediálnych dát je chránený aj po doručení ku koncovým užívateľom. V posledných rokoch sa vyvinuli a overili viaceré metódy vloženia vodoznakov do multimediálnych dát a vyčlenili sa viaceré typy systémov, ktoré techniku digitálnych vodoznakov implementujú. S vývojom techniky a systémov s vodoznakmi sa však súbežne rozvíja a zdokonaľuje aj technika útokov, ktorých cieľom je odstránenie, resp. znehodnotenie vloženého vodoznaku. Z tohto dôvodu sa pri návrhu súčasných systémov s digitálnymi vodoznakmi kladie dôraz na robustnosť, čo je požiadavkou odolnosti vodoznaku voči možným útokom.

**Keywords** — transformačná oblasť, vodoznačenie, vodoznaky, DCT, FFT, video

## I. ÚVOD

Digitálny tvar multimédií prináša so sebou veľké množstvo výhod oproti analógovej forme. Medzi tieto výhody patria hlavne jednoduchšie a spoľahlivejšie spracovávanie dát, väčšia odolnosť voči rušeniu počas prenosu a dokonalé kopírovanie, t.j. kopírovanie bez straty kvality.

Vysokorýchlostné pripojenie na Internet a digitálne multimédiá však prinášajú aj neobmedzené možnosti pre nelegálnu výrobu a distribúciu identických kopíí multimediálnych produktov. Z tohto dôvodu dochádza k jednoduchému a rýchlemu šíreniu nelegálnych dát, čo spôsobuje veľké finančné úniky pre vlastníkov autorských práv. Ochrana autorských práv a tým aj ochrana investícií a zisku sa preto stala prioritou mnohých vydavateľstiev, výrobcov a distribútorov multimediálnych produktov. Problém ochrany informačného obsahu rieši oblasť s názvom kryptografia, ktorá spočíva v utajení celého informačného obsahu multimediálnych dát. Kryptografické metódy silne chránia data počas ich prenosu, ale nedokážu chrániť dáta po ich doručení a dešifrovaní. Ďalšie možnosti na ochranu informačného obsahu multimediálnych dát poskytujú nové oblasti s názvom steganografia a digitálna vodotlač [1][2]. Obe technológie sú založené na ukrytí prídavnej informácie v obsahu multimediálnych dát a líšia sa hlavne účelmi ukrytia informácie. Hlavným cieľom steganografie je zabezpečiť utajenú komunikáciu, resp. utajený prenos prídavnej informácie na pozadí neutajenej (verejnej) komunikácie. Naproti tomu digitálna vodotlač využíva ukrytie prídavnej informácie na zabezpečenie rôznych cieľov, akými sú najmä verifikácia integrity obsahu multimédií, ochrana duševného vlastníctva, t.j. autorských práv a prenos prídavnej informácie.

## II. COMPARATÍVNA ANALÝZA Z POHLADU VKLADAJÚCICH OBLASTÍ

### A. Priestorová (časová) oblasť

Priestorová oblasť pracuje predovšetkým s audiom a videom a ich vzájomnou úpravou. Pri tejto metóde sa v prípade nevnímateľných vodoznakov využíva nedokonalosť ľudského zraku v tom zmysle, že človek nedokáže rozoznať viac ako 90 prechodov jasovej úrovne. Hlavným predstaviteľom týchto metód sú metódy vkladajúce vodoznak do bitových rovín obrazu, konkrétne do bitovej roviny zodpovedajúcej najmenej významovému bitu (*Least Significant Bit, LSB*). Daný postup spočíva v tom, že autor rovno vkladá prídavnú informáciu

do jednotlivých častí videosekvencie, čo značí, že pracuje s priestorom. Výhodou tejto metódy je, že nie je matematicky náročná a nespôsobuje zhoršenie kvality videa. Nevýhodou je menšia odolnosť voči narušeniu, prípadne odstráneniu vodoznaku.

Ďalším predstaviteľom týchto metód je technika rozprestretého spektra s využitím *DSSS (Direct Sequence Spread Spectrum)* modulácie [5]. Jej použitie nie je komplikované, avšak táto metóda sa na rozdiel od predchádzajúcej nerealizuje v statických obrazoch. Princíp spočíva vo vložení rovnakej prídavnej informácie do každej časti videosekvencie, čo tiež spôsobuje, že na jej výber by mal stačiť aj ich najmenší počet. Niektoré metódy digitálnej vodotlače na báze rozprestretého spektra reprezentujú video ako jednorozmerný diskretný signál, ktorý je získaný snímaním jednotlivých snímok videa po riadkoch, to znamená, že každý bit vodoznaku je potom vložený do 1000 alebo viac vzoriek výsledného jednorozmerného (1D) signálu a takéto systémy digitálnej vodotlače sú následne schopné dosiahnuť vysokú robustnosť voči nechceným útokom akým sú napr. filtrácia a stratová kompresia [3].

### B. Transformovaná (frekvenčná oblasť)

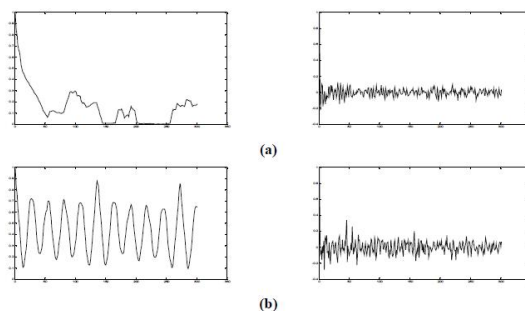
Postupy vkladania prídavnej informácie vo frekvenčnej oblasti sa aplikujú najčastejšie a mnohokrát aj s ich malými obmenami. Na začiatku sa vykoná úprava základných informácií, neskôr sa upravujú transformačné súčinitele a hľadá sa ich vzájomná podobnosť pri vložení prídavných dát. V závere sú informácie vrátené do časovej oblasti.

#### 1) DCT transformácia a jej najpoužívanejšie varianty

Diskrétna kosínusová transformácia DCT je založená na využití systému diskretných kosínusových funkcií rastúcej frekvencie s nekonštantnými zložkami vektorov transformačnej matice a úzko súvisí s diskretnou Fourierovou transformáciou DFT, pretože využíva reálnu časť transformačného jadra DFT. Podobne ako u ostatných transformácií, DCT sa snaží o dekoreláciu obrazových dát. Po dekorelácii môže byť každý transformačný koeficient nezávisle zakódovaný bez toho, aby bola znížená účinnosť kompresie. Táto časť opisuje DCT a niektoré z jej dôležitých vlastností.

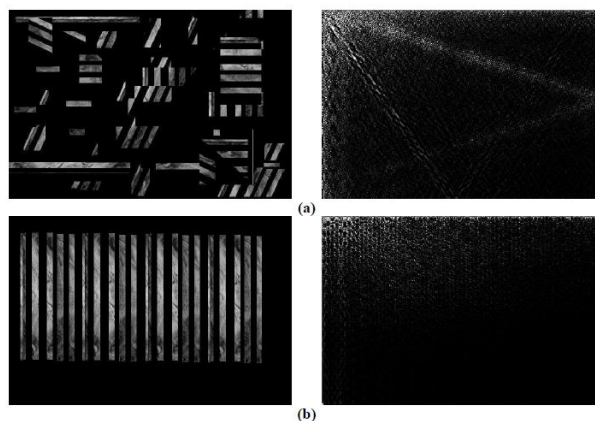
Najpodstatnejšími parametrami, ktoré charakterizujú diskretnú kosínusovú transformáciu sú dekorelácia a kompresia energie.

*Dekorelácia* - ako bolo už spomínané, principiálnou výhodou obrazovej transformácie je odstránenie redundancie medzi susednými pixelmi. To vedie k nekorelovaným transformačným koeficientom, ktoré môžu byť nezávisle na sebe zakódované. Normalizovaná autokorelácia obrazu pred a po DCT je znázornená na Obr. 1. Je zřejmé, že amplitúda autokorelácie po operácii DCT je veľmi malá vo všetkých oneskoreniach. Z tohto dôvodu, je možné odvodiť, že DCT vykazuje vynikajúce dekorelačné vlastnosti.



Obr. 1 (a) Normalizovaná autokorelácia nekorelovaného obrazu pred a po DCT; (b) Normalizovaná autokorelácia korelovaného obrazu pred a po DCT

*Kompresia energie* - výber vhodných transformačných koeficientov možno optimalizovať aj z pohľadu JPEG/MPEG kompresie, kde je známe, ktoré koeficienty budú touto kompresiou odstránené (zanedbané), a ktoré odlišné budú použité s cieľom zvýšenia robustnosti vloženého vodoznaku voči JPEG/MPEG kompresii využívajúcej dekompozíciu vstupných dát na báze DCT, čo umožňuje vložiť vodoznak v DCT oblasti a integrovať digitálnu vodotlač v statických obrazoch a video aplikovaných v reálnom čase (tento dej je ilustrovaný na Obr. 2).



Obr. 2 (a) Nekorelovaný obraz a jeho DCT; (b) Korelovaný obraz a jeho DCT

Dôležitým dodatkom teórie diskkrétnej kosínusovej transformácie je, že existuje niekoľko jej variantov, každá s čiastočnou obmenou základnej definície. Medzi najpoužívanejšie patrí jednorozmerná a dvojrozmerná DCT (1-D DCT, 2-D DCT), prípadne FFT (Fast Fourier Transform)

### C. Parametrická oblasť

Základný princíp tejto oblasti spočíva v definovaní pôvodných informácií pomocou súboru určitých hodnôt, a taktiež v tom, že pridaná informácia upravuje dané hodnoty.

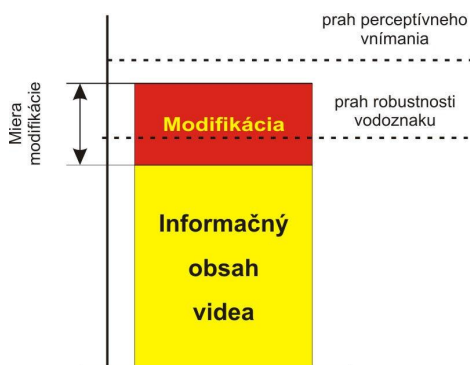
Hlavným reprezentantom metód digitálnej vodotlače s vkladaním vodoznaku v parametrickej oblasti sú metódy využívajúce fraktálové kódovanie obrazov, ktorého úlohou je obraz rozdeliť na neprekrývajúce sa bloky a každý takýto blok je vo výstupnom kóde reprezentovaný piatimi parametrami, ktorými sú pozícia najpodobnejšieho bloku ku kódovanému bloku, parameter kontrastu kódovaného bloku, parameter jasnosti kódovaného bloku a použitá izometrická transformácia.

### III. METÓDA DIGITÁLNEJ VODOTLAČE VO VIDEU NA BÁZE ČASOVEJ DIMENZIE

Metódy digitálnej vodotlače vo videu vychádzajú z rovnakej myšlienky, podľa ktorej vloženie vodoznaku do videa sa realizuje modifikáciou pôvodného videoobsahu bez výrazného zhoršenia vizuálnej kvality. Proces označenia originálneho videa vodoznakom možno formálne vyjadriť v nasledujúcom tvare:

$$V^W = V + W, \quad (1)$$

kde  $V$  je originálne video,  $V^W$  je video označené vodoznakom  $W$  a symbol  $+$  predstavuje operáciu vkladania vodoznaku, resp. modifikáciu originálneho videoobsahu. Úmyselné a neúmyselné útoky na označené video tiež spôsobujú zmenu videoobsahu, resp. miernu degradáciu kvality čo môže vyvolať výrazné poškodenie vloženého vodoznaku. Preto vkladanie vodoznaku sa obvykle realizuje modifikáciou perceptívne významných komponentov videa. Pritom miera modifikácie musí dosiahnuť určitú minimálnu hodnotu z dôvodu zabezpečenia požadovanej robustnosti vodoznaku. Na druhej strane však miera modifikácie by mala ostať pod prahom vnímania kvôli zachovaniu nedetekovateľnosti, resp. perceptívnej transparentnosti vloženého vodoznaku (Obr. 3). Väčšia modifikácia by spôsobila výraznú degradáciu vizuálnej kvality, resp. výskyt vnímateľných artefaktov vo videu.



Obr. 3 Konflikt požiadaviek na vodoznak pri modifikácii videoobsahu

Časovú dimenziu videa možno reprezentovať viacerými spôsobmi. Jeden zo spôsobov vyjadruje video vo forme vzoriek jednorozmerného (1D) signálu. Vytvorenie 1D signálu z videa je realizované snímaním videosnímkov po riadkoch (linescanning) [6]. Vzorky výsledného 1D signálu sú potom vhodne modifikované vodoznakom. Niektoré metódy digitálnej vodotlače tejto skupiny uvažujú videoobsah ako funkciu času. Proces označenia videa zohľadňuje dynamické vlastnosti videoobsahu s cieľom prispôbiť vkladanie vodoznaku k vlastnostiam ľudského zraku. V niektorých prípadoch sú vodoznaky transparentne vložené v jednotlivých snímkach. Avšak transparentnosť vodoznakov sa môže porušiť počas prehrávania videa čo sa prejavuje vo výskyte vnímateľných artefaktov pri striedaní snímkov. Metódy tejto skupiny preto vykonávajú analýzu celého videoobsahu pred samotným vkladáním vodoznaku.

#### IV. ROBUSTNOSŤ VODOZNAKOV VOČI KOLÚZNYM ÚTOKOM

Kolúzne útoky v klasickom chápaní vyžadujú znalosť viacerých kópií tých istých dát označených rôznymi vodoznakmi alebo znalosť viacerých rôznych dát označených tým istým vodoznakom. Z charakteru videa však vyplýva, že na úspešnú realizáciu kolúzneho útoku spôsobujúceho odstránenie vodoznaku postavuje samotné označené video bez potreby ďalších kópií. Najväčšiu hrozbu z úmyselných útokov na vodoznaky vo videu pravdepodobne znamenajú práve kolúzne útoky. Väčšina súčasných metód digitálnej vodotlače vo videu vykazuje nízku robustnosť voči kolúznym útokom. Kolúzny útok prvého druhu predpokladá, že obsahovo vysoko podobné, resp. rovnaké videosnímky sú označené rôznymi, t.j. nezávislými vodoznakmi. Spriemernením takýchto snímkov sa potom získava neoznačená snímka, pretože počas spriemernenia sa rôzne vodoznaky v snímkach navzájom poškodia, t.j. dochádza ku kolízii. Pri kolúznom útoku druhého druhu sa predpokladá, že všetky snímky videosekvencie sú označené tým istým vodoznakom. Na základe uvedených úvah algoritmus na zvýšenie robustnosti voči kolúznym útokom preto realizuje kompromis medzi robustnosťou vodoznaku voči obidvom druhom kolúznym útokom. Algoritmus kombinuje vkladanie toho istého vodoznaku s vkladáním rôznych vodoznakov do videa. Podstata algoritmu spočíva v segmentácii videa po scénach. V rámci jednej scény sú vkladané vždy rovnaké vodoznaky do snímkov, ale po zmene videoscény sa vkladá ďalší vodoznak. Robustnosť algoritmu voči kolúznym útokom prvého druhu je daná robustnosťou použitej metódy samotného vkladania vodoznaku voči úmyselným útokom, konkrétne voči útokom spriemernenia snímkov. Na samotné vkladanie je možné používať v podstate ľubovoľnú metódu realizujúcu vkladanie vodoznaku do statického obrazu. Podľa názoru autora je však výhodné používať metódu takú, ktorá nevyžaduje znalosť originálneho obrazu (snímky) v procese detekcie, resp. extrakcie vodoznaku, čo umožňuje vylúčiť desynchronizáciu detektora vodoznaku pri spriemernení snímkov.

#### V. ZÁVER A ZHRNUTIE

Výpočtová zložitosť súčasných metód digitálnej vodotlače vo videu na báze využitia metód vkladania vodoznakov do statických obrazov vychádza predovšetkým z potreby spracovania všetkých videosnímkov osobitne v procese vkladania vodoznaku. Výpočtová náročnosť označenia celého videa vodoznakom je potom lineárne závislá od výpočtovej zložitosti metódy použitej na samotné vkladanie vodoznaku do jednej videosnímky.

Najjednoduchším spôsobom odhadnutia výpočtovej zložitosti metód digitálnej vodotlače vo videu je meranie času, ktorý je potrebný na označenie celej videosekvencie vodoznakom. Doba trvania označenia videa vodoznakom samozrejme závisí aj od technického vybavenia, resp. výkonu zariadenia, v ktorom sa realizujú proces vkladania vodoznaku do originálneho videa.

#### REFERENCES

- [1] Cox, I.; Killian, T.; Leighton, T.; Shamon, T.: Secure spread spectrum watermarking for images, audio and video, In Proceedings of the IEEE Int.Conf. on Image Processing ICIP-96, pages 243-246, Lausanne, Switzerland, September
- [2] Zhang J., Li J., and Zhang L.: Video watermark technique in motion vector. In Proceedings of the 14th Brazilian Symposium on Computer Graphics and Image Processing, pages 179-182, October 2001.
- [3] Bodó Y., Laurent N., and Dugelay J.-L.: A comparative study of different modes of perturbation for video watermarking based on motion vectors. In Proceedings of the 12th European Signal Processing Conference, pages 1501-1504, September 2004.
- [4] Deguillaume F., Csurka G., Ruanaidh J., and Pun T.: Robust 3D DFT video watermarking. In Security and Watermarking of Multimedia Contents, volume 3657 of Proceedings of SPIE, pages 113-124, January 1999.

- [5] Kalker T., Depovere G., Haitzma J., and Maes M.: A video watermarking system for broadcast monitoring. In Security and Watermarking of Multimedia Contents, volume 3657 of Proceedings of SPIE, pages 103–112, January 1999.
- [6] Holliman M., Macy W., and Yeung M.: Robust frame-dependent video watermarking. In Security and Watermarking of Multimedia Contents II, volume 3971 of Proceedings of SPIE, pages 186–197, January 2000.



# Digitálna vodotlač vo videu

<sup>1</sup>Dušan Frkán, <sup>1</sup>Vladimír Bánoci, <sup>1</sup>Gabriel Bugár, <sup>1</sup>Dušan Levický

<sup>1</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická Univerzita Košice, Slovenská Republika

<sup>1</sup>dusan.frkan@tuke.sk, vladimir.banoci@tuke.sk, gabriel.bugar@tuke.sk,  
dusan.levicky@tuke.sk

**Abstrakt** — Cieľom tohto článku je poskytnúť základný prehľad o problematike digitálnych vodoznakov vo videu, informovať o vlastnostiach a prioritných požiadavkách, ktoré by mal vodoznak spĺňať. Zároveň pojednáva o moderných technikách vodoznačenia vo videu a taktiež o útokoch na tieto digitálne vodoznaky. Článok obsahuje aj programovú realizáciu, resp. simuláciu vodoznaku prostredníctvom vybranej techniky v prostredí MATLAB. Simulácia využíva metódu vkladania vodoznakov v priestorovej oblasti použitím LSB metódy s následnou analýzou dosiahnutých výsledkov.

**Kľúčové slová** — Vodoznak, útok, video, simulácia vodoznaku, techniky vodoznačenia, priestorová oblasť, LSB

## I. ÚVOD DO PROBLEMATIKY VODOZNAKOV

Súčasný rozmach v používaní vysokorýchlostných počítačových sietí, Internetu, či World Wide Webu prinieslo revolúciu v spôsobe, akým môžu byť digitálne dáta distribuované. Analógová forma multimédií bola nahradená ich digitálnou podobou prakticky vo všetkých oblastiach ľudského života. Medzi hlavné výhody digitálnej multimediálnej formy patrí predovšetkým ľahké spracovanie a ukladanie, kompresia, či lepšia odolnosť voči šumu. Avšak, výhody digitálnej revolúcie neboli dosiahnuté bez nevýhod, ktoré predstavujú rozšírené a ľahko dostupné prístupy k multimediálnemu obsahu, či možnosti vytvorenia neobmedzeného množstva kópií bez veľkej námahy. Uvedené aspekty viedli čoraz častejšie k myšlienke vytvoriť ochranu autorských práv a práve digitálne vodoznaky reprezentujú technológiu, ktorá ma predpoklady slúžiť k tomuto účelu.

Vodoznak predstavuje digitálnu informáciu alebo údaj vložený v akomkoľvek multimediálnom objekte. Hlavným cieľom autora digitálneho vodoznaku je vložiť takúto prídavnú informáciu nepozorovane zvyčajne do redundantných častí dát, ideálne bez zásadnej degradácie kvality pôvodného obrazu, či videa a s dostatočnou odolnosťou voči úmyselným útokom. Podmienkou je možnosť jeho detekcie alebo extrakcie iba vlastníkom, čo výrazne uľahčuje určiť pôvodcu prvotných dát.

## II. DIGITÁLNA VODOTLAČ A OCHRANA OBSAHU MULTIMÉDIÍ

Cieľom ukrytia prídavnej informácie (information hiding), resp. vloženia prídavnej informácie do obsahu multimédií, je vymedzenie dvoch veľmi príbuzných oblastí manipulácie s obsahom multimédií a tými sú:

- steganografia,
- digitálna vodotlač (digital watermarking)

**Steganografia** (steganography) je veda, ktorá sa zaoberá skrývaním dát alebo správ do digitálnych súborov a iných digitálnych štruktúr. Na rozdiel od kryptografie, kde je hlavným cieľom kódovanie správy, úlohou steganografie je skryť skutočnosť, že správa je odoslaná. Dôraz kladie na utajený prenos informácie, ktorý je realizovaný na pozadí neutajenej komunikácie[1][2]

**Digitálna vodotlač** (digital watermarking) predstavuje techniku, ktorá poskytuje riešenie dlhoročných problémov s ochranou autorských práv digitálnych dát. Digitálne vodoznaky predstavujú určité časti informácie pridané do digitálnych dát (audio, video alebo statické

snímky), ktoré môžu byť neskôr objavené alebo extrahované. Tieto informácie môžu byť buď vo forme textových údajov alebo ich môže predstavovať samotný obrázok. Digitálne vodoznaky zostávajú neporušené aj pri manipuláciách s nimi (úprava alebo prenos vodoznaku), čo umožňuje chrániť autorove vlastnícke práva v digitálnej podobe.

#### A. Vloženie vodoznaku a jeho extrakcia

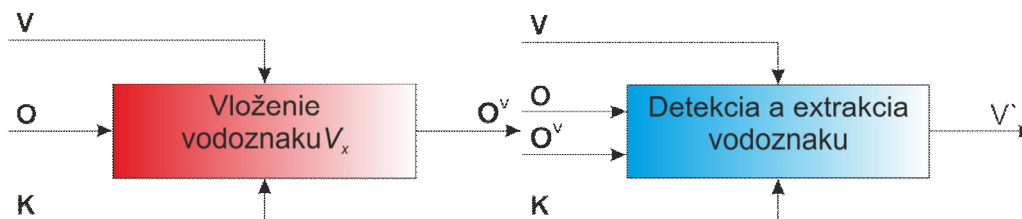
Systém digitálnej vodotlače pozostáva z dvoch základných blokov, a to z bloku vloženia a extrakcie vodoznaku (watermark embedder) zobrazeného na Obr. 1.

Tieto bloky sú definované pomocou nasledujúcich parametrov  $[O, O^V, V, K, V_x, E_x, V]$ , pričom  $O$  predstavuje originálne dáta (obraz),  $O^V$  definuje vodoznakom označené dáta (obraz),  $V$  je prídavná informácia a  $K$  označuje aplikované kľúče. Vložením prídavnej informácie  $V$  do pôvodných dát  $O$  pomocou tajného kľúča  $K$  sa dosiahne lepšia celková ochrana pri práci s vodoznakom. Platí, že aplikovaním väčšieho množstva kľúčov sa zvyšuje šanca, že samotný vodoznak nebude extrahovaný.

Algoritmus používaný pre vloženie prídavnej informácie definuje vzorec (1) a extrakciu vzorec (2):

$$V_x: O \times V \times K \longrightarrow O^V \quad (1)$$

$$E_x: O^V \times K \longrightarrow V \quad (2)$$



Obr. 1 Schéma pre vloženie a extrakciu vodoznaku

#### B. Primárne požiadavky kladené na metódy digitálnej vodotlače

Každý digitálny vodoznak musí spĺňať vysoké kritériá na základe svojho určenia. Tieto požiadavky sú rozdelené na 5 základných parametrov:

- robustnosť, resp. odolnosť (robustness),
- bezpečnosť (security),
- perceptívna transparentnosť (perceptual transparency),
- nedetekovateľnosť (undetectability),
- kapacita (capacity),
- možnosť viacnásobného vloženia vodoznakov (multiple embedding).

#### C. Hlavné rozdelenie digitálnych vodoznakov

Rozdelenie vodoznakov podľa viditeľnosti definuje nasledovné [3]:

- *viditeľný vodoznak*, ktorý je na pohľad postrehnuteľný. Je potrebné, aby boli splnené dve základné požiadavky a to nemožnosť odstránenia prídavnej informácie neautorizovanou osobou bez narušenia pôvodných dátových informácií, a taktiež schopnosť odolať pokusom o falzifikáciu.
- *nepostrehnuteľný vodoznak* predstavuje dodatočné dáta, ktoré sú zakomponované v originále a je možné ich nájsť iba špeciálnymi metódami. Input vodoznaku zabezpečuje povolaná osoba tajným kľúčom, čím zároveň predstavuje prekážku pri pokusoch o jeho odstránenie. Dôležitým parametrom pre nepostrehnuteľné vodoznaky je odolnosť a nevnímateľnosť.
- *duálny vodoznak* – táto skupina predstavuje spojenie postrehnuteľných a nepostrehnuteľných vodoznakov v originálnych dátach. Princíp spočíva v ich postupnom vkladaní.

Na základe krehkosti vodoznakov poznáme :

- *krehké vodoznaky* sa vyznačujú redukovanou odolnosťou. Ich využitie spočíva

- v hľadání poškodenia alebo obmeny vodoznaku, najmä zmeny v ich zdrojových informáciách.
- *polokrehké vodoznaky* tvoria samostatný celok krehkých vodoznakov. Výhodou polokrehkých vodoznakov je schopnosť rozlíšiteľnosti medzi overenou a neoverenou úpravou informácií ako napr. A/D a D/A konverzia, dekvantizácia alebo gamma korekcia.
  - *robustné vodoznaky* musia byť odolné voči rozličným modifikáciám označených informácií.

#### D. Útoky na digitálne vodoznaky

Úlohou útokov je narušiť alebo zničiť digitálny vodoznak nachádzajúci sa v originálnych dátach. Útok činí úmyselný, systematický a usmerňovaný dej uskutočňovaný neautorizovanými osobami, tzv. útočníkmi. Rozbor účinku týchto útokov je podstatným faktorom pri navrhovaní digitálneho watermarkingu. Útoky delíme na 5 primárnych okruhov [4]:

- útoky odstraňujúce vodoznak,
- kryptografické útoky,
- protokolové útoky,
- geometrické útoky.

Cieľom útokov odstraňujúcich vodoznak je zlikvidovať určitý vodoznak z označených dát bez toho, aby bola prelomená ochrana použitého algoritmu, kľúča, ktorého úlohou je vkladať tento vodoznak. Podmienkou takýchto útokov nemusí byť vždy len úplné odstránenie vodoznaku. Postačujúcim je aj jeho podstatné narušenie alebo aj znemožnenie jeho lokalizácie. Súčasťou tohto okruhu útokov sú napríklad kvantizácia, filtrácia, kolúzne útoky, útok kopírovaním, atď.

V kryptografických útokoch možno nájsť určitú symboliku s útokmi používanými v oblasti kryptografie. Úlohou je nájsť cestu, ktorou by mohol byť vložený vodoznak odstránený, resp. vložený ďalší, nepravý. V odbornej literatúre sa vyskytujú pod názvom *cryptographic attacks*.

V praxi sa využívajú tieto konkrétne metódy útokov:

- metóda totálnych skúšok,
- metóda hrubého násilia

*Metóda totálnych skúšok* spočíva v skúšaní každého dostupného kľúča. Primárnym parametrom tejto metódy je dostatočný časový úsek, určený na preskúmanie každej metódy. Závisí od dĺžky kľúča.

*Metóda hrubého násilia* - hlavnou súčasťou tejto metódy je nutnosť použitia detektora a informácií, na ktoré bol použitý vodoznak. Detektor slúži na zachytenie zmeny informácie. Postup spočíva v malých obmenách týchto dát, pričom sa na výstupe zachytáva úroveň citlivosti detektora a následne porovnáva s ostatnými hodnotami pri daných zmenách.

Primárnou úlohou *protokolových útokov* je poškodiť podstatné zákonitosti, využívané pri vodoznakoch. Najznámejšími typmi takýchto útokov sú technika invertibilných systémov a útok kopírovaním.

*Technika invertibilných systémov* využíva vloženie pridanej informácie do originálnych dát pomocou inverzných funkcií, čiže opačných od tých, ktoré použil autor. Vo výsledných dátach sa preto budú nachádzať dva vodoznaky, originálny od pôvodného autora a podvrh, napodobenina útočníka. Kvôli tomu sa v súčasnosti kladie dôraz na neinvertibilitu vodoznaku. Riešením tohto problému je použitie tzv. jednocestných funkcií.

*Útok kopírovaním* sa používa, keď je na identifikáciu odlišných dát využitý rovnaký vodoznak. Princíp spočíva vo vybratí vodoznaku z označených dát a jeho následnom prekopírovaní do výsledných dát. Hodnotenie výstupu je podobné ako v metóde kolúzných útokov.

Cieľom *geometrických útokov* nie je zlikvidovať vodoznak z originálnych informácií. Pri strate synchronizačnej závislosti medzi časťou pre vkladanie resp. extrakciu vodoznaku a dátami, dochádza k značnému poškodeniu vodoznaku, ktorý bol vložený do dát. Táto oblasť zahŕňa predovšetkým otáčanie obrazu, jeho redukciu resp. násobenie. Ich prioritou je znemožnenie nájdania pridanej informácie v originálnych dátach.

### III. SPÔSOBY VLOŽENIA VODOZNAKU

Spôsoby vloženia vodoznaku je možné adekvátne zaradiť do niekoľkých skupín, podľa toho, v akom priestore sa daný vodoznak bude nachádzať. Z tohto dôvodu možno rozlíšiť metódy digitálnej vodotlače s vkladáním vodoznaku v [5][6]:

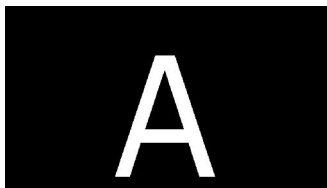
- priestorovej oblasti,
- transformovanej oblasti,
- parametrickej oblasti.

Tieto spôsoby uskutočňujú vloženie prídavnej informácie do originálnych dát s cieľom zvýšenia ochrany, najmä odolnosti, pred neautorizovanými osobami a preto sa vyžaduje ich aplikácia do väčšiny použitých informácií. Nevýhodou je, že klesá kvalita obrazov alebo videí, ak je pridaná informácia prítomná v tých najvýznamnejších úsekoch.

Definované oblasti sú založené na nepresnostiach samotného človeka, resp. jeho sluchového a zrakového systému. Cieľom je však vytvoriť rovnováhu medzi týmito nežiaducimi a potrebnými vlastnosťami vzniknutými pri aplikácii vodoznaku práve pomocou takýchto systémov.

### IV. SIMULÁCIA METÓDY LSB

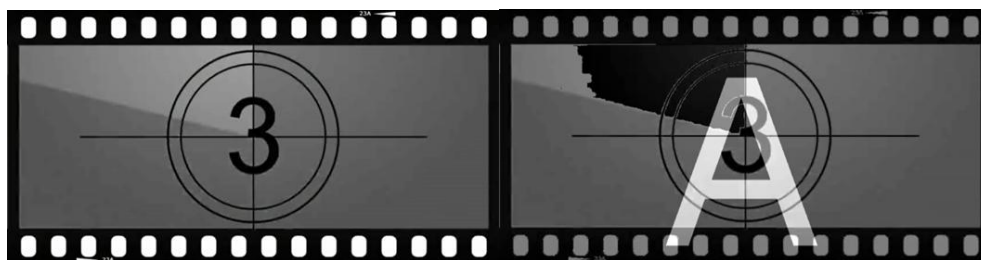
Simulácia vodotlače prostredníctvom metódy LSB vo videu má za úlohu demonštrovať zmysel používania digitálnych vodoznakov vzhľadom na dosiahnutie požadovaných parametrov kladených na vodoznaky. Metóda LSB (Least Significant Bit), teda najmenej významového bitu je charakterizovaná priamym vložením prídavnej informácie do videa a vyznačuje sa malou náročnosťou a taktiež jej aplikácia nepreukazuje výraznú degradáciu kvality. Dokáže prežiť jednoduché útoky, ako je orezanie alebo pridanie šumu [7]. Priame vkladanie vodoznaku sa realizuje v priestorovej oblasti použitím dvojrozmerného poľa obrazových prvkov. Vychádzajúc z faktu, že ľudské oči nie sú veľmi prispôbené na minimálne farebné rozdiely, spracovanie týchto rozdielov v LSB metóde nebude viditeľné



voľným okom.

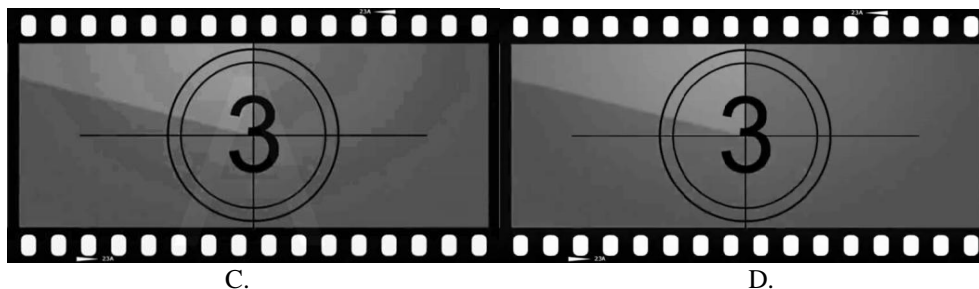
Obr. 2 Aplikovaný vodoznak

Aplikovaný vodoznak má teda rozmery 854 x 440 pixelov, ktoré sú úplne zhodné s daným videom. Kým hodnota 854 pixelov udáva šírku vodoznaku, resp. videa, tak hodnota 440 pixelom udáva výšku vodoznaku a teda aj videa. Je známy parameter videa, ktorý udáva, že frekvencia snímok je stanovená na 25 snímok za sekundu. Keďže je známa aj celková dĺžka trvania videa, tak jednoduchým matematickým výpočtom vieme určiť, že celkový počet snímok bude rovný 125. Video predstavuje útvar (objekt), v ktorom snímky nasledujú kontinuálne za sebou. Môžeme ho teda popísať maticou so štyrmi rozmermi v tvare  $a=854 \times 440 \times 3 \times 125$ , pričom bolo už uvedené, že 854, resp. 440 pixelov udáva šírku, resp. výšku videa, 3 udáva, koľko zložiek RGB modelu dané video obsahuje a 125 identifikuje z koľkých snímok sa video skladá.



A.

B.



Obr. 3 Originál snímka bez vodoznaku (A.), označená snímka s vloženým vodoznakom: v 1. rovine (B.), 5.rovine (C.), 8.rovine (D.)

V závere simulácie bolo po aplikovaní roviny vodoznaku miesto vybranej bitovej roviny každej snímky vyrobené nové zrekonštruované video obsahujúce tieto novovytvorené upravené snímky. Ôsma bitová rovina je najmenej významná, čo spôsobuje aj jej podstatný výskyt šumu. Čím vkladáme vodoznak do významnejšej bitovej roviny, tým je viditeľnosť vodoznaku vyššia ako ukazuje Obr.3 (B.), kde Obr.3 (D.) práve dokazuje, že vodoznak nachádzajúci sa v poslednej a teda najmenej významovej bitovej rovine každej snímky je voľným okom neviditeľný a takmer identický s originálnou snímkou videa.

V Tab. 4 je zaznamenané porovnanie hodnôt PSNR vybraných použitých bitových rovín, pričom parameter bit per pixel (*bpp*) bol pre všetky roviny rovnaký s hodnotou  $b = 0.0693$ .

Table 1.  
Kvalitatívne porovnanie obrazovej kvality snímky vzhľadom na využitú bitovú rovinu

Bitová rovina	PSNR [db]
I.	15,2656
IV.	34,6769
VI.	40,8934
VIII.	52,2656

## V. CONCLUSION

Z uvedených obrázkoch vidno, že vodoznak vložený v prvej rovine je nadmieru viditeľný a teda jeho funkcia ochrany obsahu je nespĺnená. Ôsma bitová rovina je najmenej významná, čo spôsobuje aj jej podstatný výskyt šumu. Čím vkladáme vodoznak do významnejšej bitovej roviny, tým je viditeľnosť vodoznaku väčšia. Obr. 3 dokazuje, že vodoznak nachádzajúci sa v ôsmej a teda najmenej významovej bitovej rovine každej snímky je voľným okom neviditeľný a takmer identický s originálnou snímkou videa. Je potrebné poznamenať, že hlavnými nevýhodami použitej metódy LSB je jej menšia odolnosť voči narušeniu, prípadne odstráneniu vodoznaku. Napríklad útok v podobe stratovej kompresii dokáže úplne zničiť vodoznak. Ešte účinnejším útokom je však nastavenie všetkých LSB bitov na hodnotu 1, čím sa vodoznak stáva viditeľným a tým je pre útočníkov viac ako ľahké ho odstrániť.

## POĎAKOVANIE

The preferred spelling of the word “acknowledgment” in American English is without an “e” after the “g.” Use the singular heading even if you have many acknowledgments. Avoid expressions such as “One of us (S.B.A.) would like to thank ... .” Instead, write “F. A. Author thanks ... .”

## REFERENCES

- [1] Pope, Michael Brian: Digital steganography. Department of Management & Information Systems, Mississippi State University, 2012.
- [2] Celik, M. U., Sharma, G., Tekalp, A. M., Saber, E.: Reversible Data Hiding. in Proceedings of International Conference on Image Processing, Vol. 2, pp. 157-160, September 2002.
- [3] Li, Chang-Tsun: Multimedia Forensics and Security. University of Warwick, 2008.
- [4] Hong Zhao: Watermark attacks. University of Maryland, 2002.
- [5] Chan P.-W.: Digital Video Watermarking Techniques for Secure Multimedia creation and Delivery. PhD thesis, The Chinese University of Hong Kong, 2004.
- [6] Weijia Jia, Fung Po Tso, Zhen Ling, Xinwen Fu: Blind Detection of Spread Spectrum Flow Watermarks. City University of Hong Kong, 2009.
- [7] Foriš, Peter: Perceptívne metódy digitálnej vodotlače v statických obrazoch. Písomná práca k dizertačnej skúške. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2006.

# Pattern recognition in motion with sensors of mobile devices

<sup>1</sup>*Patrik Mucha*, <sup>2</sup>*Miroslav Biňas*, <sup>3</sup>*Marek Novák*

<sup>1,2,3</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>patrik.mucha@student.tuke.sk, <sup>2</sup>miroslav.binas@tuke.sk, <sup>3</sup>marek.novak@tuke.sk

**Abstract** – In this paper We present a model for human activity recognition from accelerometer and orientation sensor on an Android mobile device. Our model can recognize three human activities: sitting, staying, walking and a state when mobile phone is lying on a table. The recognition is based on a method similar to decision tree. This method uses mean, variance and vector magnitude of an accelerometer vector. The experiments were realized on an Android application with 10 respondents with promising accuracy above 70% on various models of devices.

**Keywords** – Accelerometer, activity recognition, android, orientation sensor

## I. INTRODUCTION

Research and development in the area of human activity and gait recognition is highly actual nowadays. It is fostered by mounting various sensors into new smart phones. Since the smart phones have these sensors built-in there is a fruitful progress in this research area. However many projects have been done, the problem of motion recognition using wearable sensors is relatively still in the beginning.

The majority of currently available smart phones on a market contain tri-axial accelerometers that can measure acceleration in all three dimensions. These accelerometers are also capable of detecting the orientation of the device (helped by the fact that they can detect the direction of Earth's gravity), which can provide useful information for activity recognition. Accelerometers were initially included in these devices to support advanced game play and to enable automatic screen rotation but they clearly have many other applications. We present an algorithm for detection four different actions of a user: staying, walking, sitting or lying (the position and orientation of a device is nearly the same) and state when mobile phone is put on a table. For action recognition We have created a model similar to decision tree based on a features as: mean, variance of accelerometer vector and vector magnitude.

Section 3. presents the analysis of human motion detection and recognition. Section 4. presents our model and algorithms used, experiments performed and results.

One of the goals of this paper is to find out whether it is possible to recognize human activities using sensors of mobile devices and then design and implement system for human activities recognition. Another goal of this paper is to design and implement feature for person identification using accelerometer sensor's data.

## II. PATTERN RECOGNITION ANALYSIS

Motion is usual in everyday life of human beings. The most common method of motion is walking. By watching on and monitoring human gait can be seen that many people have different style of walking, like unique motion pattern. This was one of the inspirations for this thesis.

### A. What is gait

Gait is repeating sequence of steps where one step is motion of one leg forward and then motion of the other leg forward to default position. When watching on gait in the beginning of each step the leg accelerates and in a certain point the leg begins to decelerate. The same scenario repeats even in the case of the second leg.

It can be monitored gait velocity and the length of step. These features can be labeled as the motion pattern of walking. The problem of the gait analysis can be divided into two parts: biometric gait identification and gait identification for biomedical purposes. [1]

### B. How to register and identify motion and gait

The existent gait identification methods can be grouped into three categories: machine vision (MV) based, floor sensor (FS) based and wearable sensor (WS) based. By MV-based gait recognition, gait is captured using a video camera from the distance. Video and image processing techniques are then employed to extract gait features for recognition purposes. By FS-based gait recognition, a set of force plates are installed on the floor. Such force sensors measure gait related features, when a person walks over them. By WS-based gait recognition the information is collected using body worn motion recording sensors, mostly accelerometers. Sensors can be located at different positions on the observed person, such as pocket, leg, waist, belt, hip, etc. The acquired acceleration signal of the gait is then used for motion and gait identification. [1]

For motion recognition was used a mobile device - smart phone. Many smart phones have various sensors built-in including accelerometer.

## III. SYSTEM DESIGN

### A. Sensors data

In our project we have used data provided by accelerometer and orientation sensor of smart phone for human activity recognition. An acceleration sensor in a smart phone measures the acceleration applied to the device, including the force of gravity [2]. Acceleration of a device is formulated as (1), where  $\sum F$  is the sum of forces applied to the sensor itself,  $m$  is phone's mass and  $g$  is gravity acceleration of a magnitude  $9,81m/s^2$ . Since gravity is included, the accelerometer reads value 9,81 in  $z$  axis, when phone is lying on a table.

$$A = -g - \sum(F/m) \quad (1)$$

The acceleration force is measured in all three axis in 3D coordinate system. Since the user can put a smart phone in his/her pocket and its orientation can vary, it is necessary to calculate acceleration vector magnitude (using formula 2) which is orientation independent. The same approach was used in various projects, i.e. [3] [4] [5].

$$v = \sqrt{x^2 + y^2 + z^2} \quad (2)$$

### B. States and Activity recognition

Smart phone can be in two states:

- 1) in *stationary state*, or
- 2) in *motion*.

To detect motion an arithmetic mean and a variance of an acceleration vector are calculated. If the variance value is lower than fixed constant the smart phone is in stationary state or is moving at constant speed. Possible human activities are: *staying*, *sitting (lying)* or smart phone *put on a table*. These activities can occur while driving a car or sitting in a bus when the vehicle moves at constant speed too. To differentiate these activities the data from orientation sensor is used. When the Y-axis and Z-axis values of orientation vector are very close to zero then the smart phone is put on the table. When long side of the smart phone is heading to the floor or sky then the user is sitting or lying on a bed. Sitting means orientation of the smart phone rotated along the long side of device while monitoring Z-axis of orientation vector in fixed range. Any other orientation of the smart phone can be considered as staying.

If the variance is higher than above mentioned constant then the smart phone is accelerating or decelerating. Possible human activity is walking. Any other motions like shaking or jumping are undefined in our system and are labeled as OTHER.

### C. Detection of Walking

When the motion is detected another algorithms are applied. The accelerometer measures raw data and the vector magnitude is calculated in every moment.

The first step in walking detection algorithm is application of low-pass filter on acquired acceleration data. On every magnitude value the low-pass filter is applied to smooth these values. After filter application the new values are obtained. Using these new values it is easier to design an algorithm to walking recognition.

Various projects defined walking as a part of a graph from one peak to following peak [6] [7] [8]. Project [9] also describes advantages of the peak detection algorithm in process of walking detection. Based on above mentioned facts we chose to implement similar algorithm in our work and it is next step in our walking detection algorithm. Our algorithm finds the local minimum

and local maximum values in list of vector magnitudes after filter application. In contrast with [10] where derivative is used to find peaks we use a bit simpler algorithm. In our peak detection algorithm we calculate in every moment two values:

- 1) difference between the previous and current acceleration vector magnitude, and
- 2) difference between current and next acceleration vector magnitude.

If the value change from negative to positive, it indicates local maximum. Every peak comes with the timestamp and after all peaks in time period are detected the time differences between every two neighboring are calculated. If all of obtained differences are higher than fixed constant then the user is walking.

#### D. Identification of a Person

In our system we also implemented the algorithm for identification of the person, say - owner of the smart phone. In the application for mobile device the pattern specific for walking is stored. The data for pattern creation was measured by smart phone *Samsung Galaxy S3* international version at fastest rate, analyzed on a computer and finally stored in application. In present, this algorithm is optimized only for above mentioned device. When the mobile application detects walking for at least 10 seconds in every specific time period, say every minute, the algorithm compares the current magnitude values with the values of stored pattern. First step in this algorithm is to sync these two patterns and then calculate the difference between each value from two patterns. Result will be average difference between mentioned patterns.

### IV. EXPERIMENTS AND THEIR RESULTS

For the motion and human activity recognition we have conducted measurements with 10 users, where each user repeated each activity 2 times. The first measurement was accomplished with 1 second interval and the other one with 2 seconds interval of human activities recognition. Altogether 80 measurements were collected and analyzed. The monitored users were young people in the age of 20 till 30. The users wore a smart phone in their pocket as and were asked to perform walking for 30 seconds and every other activity for 60 seconds. After performing an activity We compared time from a counter to real time the user was performing an activity.

For the person identification we have conducted measurements with 5 users. The monitored users were young people in the age of 20 till 30 and they wore a smart phone as is mentioned above. Every user used in this test only one smart phone, the same device, which was *Samsung Galaxy S3*. They were asked to walk for at least 1 minute. The tests were realized at least 10 times in various days.

The summary results are presented in table 1 and table 2. This table is in a form of confusion matrix where columns of the matrix represent the instances in a predicted class and rows represent the instances in an actual class. The values are number of seconds how the real activity was recognized. For example if a user was walking for 60 seconds, but application recognized 54 seconds of walking and 6 seconds of staying, the first row would be 54, 0, 0, 0, 6.

Table 1  
 Accuracy of human activities recognition - 1 second interval

	<b>Walking</b>	<b>Sitting</b>	<b>Staying</b>	<b>Table</b>	<b>Other (undefined)</b>
Walking	<b>213</b>	0	0	0	87
Sitting	2	<b>566</b>	21	0	11
Staying	3	15	<b>579</b>	1	2
Table	0	0	19	<b>581</b>	0
Accuracy (%)	<b>71.00</b>	<b>94.33</b>	<b>96.50</b>	<b>96.83</b>	

Table 2  
 Accuracy of human activities recognition - 2 seconds interval

	<b>Walking</b>	<b>Sitting</b>	<b>Staying</b>	<b>Table</b>	<b>Other (undefined)</b>
Walking	<b>248</b>	0	0	0	52
Sitting	4	<b>570</b>	18	0	8
Staying	0	10	<b>584</b>	4	2
Table	0	0	24	<b>576</b>	0
Accuracy (%)	<b>82.67</b>	<b>95.00</b>	<b>97.33</b>	<b>96.00</b>	

The acquired results in table 1 and table 2 show that it is possible to detect motion of human and recognize current user activity. Results show that with service working in 2 seconds recognize



interval the better results are obtained. Walking detection accuracy is not so high because every manufacturer built in different sensors in their smart phones which have various sensitivity and magnitudes and the algorithms were implemented to work on various smart phones.

The results of person identification are presented in table 3. In the application the walking pattern of User 1 is stored. In each row are three values:

- mean difference between walking patterns of User 1 and User X
- lowest difference between walking patterns of User 1 and User X
- highest difference between walking patterns of User 1 and User X

Table 3  
Differences in pattern relative to User 1

User	Mean	Lowest	Highest
User 1	1.97991	1.38491	3.17588
User 2	2.31173	1.62212	3.01258
User 3	2.40192	1.84215	3.16117
User 4	2.26137	1.49375	3.43718
User 5	2.29123	1.62241	3.14623

Based on the results in table 3 it is not possible to clearly identify the person. The results show that there are high differences between each measurement. The algorithm for person identification compares the current walking pattern and the stored walking pattern. This walking pattern is stored in the application and was created from analyzed accelerometer's data on computer and never changed what can cause these bad results. In the present the new self-learning algorithm is designing and implementing. It will monitor the differences in pattern in fixed intervals and will modify stored pattern what can lead to better results.

## V. CONCLUSION

In this paper We have presented an algorithm for human activity recognition on a mobile device utilizing data from accelerometer sensor. By this algorithm we are able to recognize four different states: a user is walking, staying, sitting or the smart phone is lying on a table. The recognition model is based on a method similar to decision tree.

Experiments for human activity recognition were realized by an application on HTC, Sony and Samsung smart phones with 10 users at all and experiments for person identification realized 5 people using only one and the same smart phone which was Samsung Galaxy S3 International version.

Currently we are using activity recognition for a long term human behavior monitoring for the purpose of anomaly detection in their daily habits what can help in identification of a person.

## ACKNOWLEDGMENT

This work is the result of the project implementation: *Development of the Center of Information and Communication Technologies for Knowledge Systems* (ITMS project code: 26220120030) supported by the *Research & Development Operational Program* funded by the *ERDF*.

## REFERENCES

- [1] D. Z. Sebastijan Sprager, "A cumulant-based method for gait identification using accelerometer data with principal component analysis and support vector machine," *WSEAS TRANSACTIONS on SIGNAL PROCESSING*, vol. 5, pp. 369–378, 2009.
- [2] "Android developer page: Motion sensors." [Online]. Available: <http://developer.android.com/guide/topics/sensors/sensorsmotion.html#sensors-motion-accel>
- [3] B. L. B. G. S. L. Lin Sun, Daqing Zhang, "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations," *Ubiquitous Intelligence and Computing*, pp. 548–562, 2010.
- [4] B. P. M. M. Sauvik Das, LaToya Green, "Detecting user activities using the accelerometer on android smartphones," (Visited on 01/2013). [Online]. Available: [http://www.truststc.org/reu/10/Reports/DasGreenPerezMurphy\\_Paper.pdf](http://www.truststc.org/reu/10/Reports/DasGreenPerezMurphy_Paper.pdf)
- [5] J. R. Pekka Siirtola, "Recognizing human activities user independently on smartphones based on accelerometer data," *International Journal of Artificial Intelligence and Interactive Multimedia*, vol. 1, no. 5, pp. 38–45, 2012.
- [6] J. Scarlett, "Enhancing the performance of pedometers using a single accelerometer," (Visited on 12/2012). [Online]. Available: [http://www.analog.com/static/importedfiles/application\\_notes/47076299220991-AN\\_900.pdf](http://www.analog.com/static/importedfiles/application_notes/47076299220991-AN_900.pdf)
- [7] S. H. R. v. L. D. R. Mirko Brandes, Wiebren Zijlstra, "Accelerometry based assessment of gait parameters in children," *Gait & Posture*, vol. 24, pp. 482–486, 2006.
- [8] P. S. T. K. Melis Oner, Jefry A. Pulcifer-Stump, "Towards the run and walk activity classification through step detection - an android application," *Conf Proc IEEE Eng Med Biol Soc*, 2012.
- [9] M. W. Felix Scholkmann, Jens Boss, "An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals," *Algorithms*, pp. 588–603, 2012.
- [10] R. Libby, "A simple method for reliable footstep detection on embedded sensor platforms," 2008, (Visited on 01/2013). [Online]. Available: <http://ubicomp.cs.washington.edu/uwar/libbypeakdetection.pdf>

# Global electrical power planning, investments and projects

<sup>1</sup>Vladimír GÁLL, <sup>2</sup>Alexander MÉSZÁROS

<sup>1,2</sup>Department of Electric Power engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>vladimir.gall@student.tuke.sk, <sup>2</sup>alexander.meszaros@tuke.sk

**Abstract**— that work deals with investments and projects in electric power industry. There are parts about the value of electric energy, the efficiency of generating power plant and global energy scene.

**Keywords**— electric energy, electrical power industry, world economy

## I. THE VALUE OF ELECTRIC ENERGY

Electric energy is versatile, universal, clean to use, easy to distribute and supreme to control. That is important, it is now established that electric energy has better productivity in many applications in factories than most other energy forms [1].

All this leads to the wider utilisation of electric energy and its replacement of other forms of energy in many practical uses. Demand for electric energy is growing globally at a rate higher than that of economic growth and in many countries, at almost two times higher that of demand for primary energy sources. With the type of technologies and applications in factories that already exist and there is nothing to stop advancement of electric energy. That is assumed a higher share of the electric energy market. Saturation of electric energy use isn't yet in sight, even in advanced economies where production of electric energy claims more than half of the primary energy which is used in factories. For other sector than for the sector of transport, electric energy can satisfy most human energy requirements. It is assumed that, by the middle of the 21st century, almost 70 percent of energy needs in some industrialised countries will be satisfied by electric energy [2]. Electric energy has become an important ingredient in human life and it is essential for modern living and business. Its interruption can incur large losses and create large havoc in major cities and urban centres and its disruption, even if transient may cause tremendous inconvenience. That is why, continuity of electric energy supply is significant. Also, with the widespread use of computers and other voltage and frequency sensitive electronic equipment, the importance of the quality of supply electric energy has become evident. A significant proportion of investment, in the electricity supply industry, goes into the reserve generating plant, standby equipment and other redundant facilities needed to ensure the continuity and high quality of the supply. Optimisation of investment in the electricity supply industry requires understanding of markets, prediction of future demand and an approach based on integrated resource planning.

## II. ENVIRONMENTAL CONCERNS AND THE EFFICIENCY OF GENERATING PLANT

Recently, there has been growing environmental concern about the role of energy, particularly electric energy production, in causing emissions that have local, regional and possibly global effects, with long-term detrimental implications. This has prompted emphasis on efficiency as a very important way of curbing emissions. Generation of electric energy isn't an efficient conversion process. The large part of the existing inventory of generating plant, globally, has a gross generation efficiency of less than 33 percent. The classic coal firing power stations have an efficiency that averages 25 percent. Significant improvements in efficiency have been achieved during recent years, almost to the range of 0.2 percent per year. Gross global generation efficiency, over 33 percent now, was estimated to be around 31 percent in the 1980s, 29 percent in the 1970s and as low as 27 percent in the 1960s. Modern, large and efficient thermal generating facilities utilising pulverised coal have an efficiency approximating

between 43 percent and 44 percent. Modern combined cycle gas turbine plants claim an efficiency nearing 60 percent, and in the relatively few instances of combined heat and power arrangements, efficiencies of fuel utilization can exceed 80 percent. However, owing to the very large inventory of existing inefficient generating facilities worldwide, the average efficiency of generation of electric energy won't significantly improve in the future. When the system losses of 10 percent are in industrialised countries, and more than 25 percent in many developing countries, are taken into account. Then the amount of energy reaching consumers in the form of electric energy is less than 30 percent of the calorific total of the fuel infeed. In some developing countries it may be even lower than 20 percent.

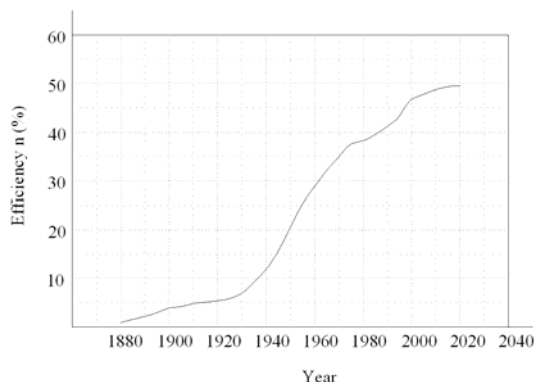


Fig. 1 Development of efficiency generating electric energy in thermal plants

### III. ELECTRIC ENERGY AND THE WORLD ECONOMY

There has always been a tight coupling between energy use and economic growth. Demand for electric energy has always displayed good correlation with that of global economic growth, as measured by gross national product (GNP) in purchasing power parity dollars. During the past 25 years, electric energy growth averaged 4 percent per annum, which is a high rate of growth. In the future, electric energy demand growth is expected to match the growth of the world economy. This is expected to average around 2.5 percent to 3.5 percent annually during the next few years. The International Energy Agency (IEA) and the International Atomic Energy Agency (IAEA) estimate that global electric energy production will increase at an annual average rate of 2.7 percent to 3 percent during the first decade of the 21st century [19]. Therefore it is expected that total electric energy production in 2010 will amount to around 20000 TWh and in 2020 to 25880 TWh. Most of this growth is going to occur in developing countries, particularly in south-east Asia, a region that is enjoying a rapid economic growth. In 2030 global electric energy production is expected to exceed 28000 TWh. One Half of this amount will be accounted for by developing countries [20].

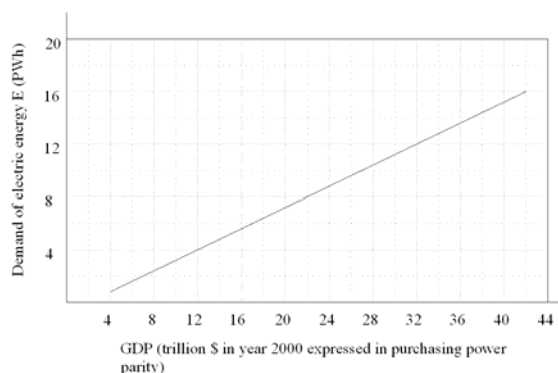


Fig. 2 Electricity demand as a function of world gross national product – GNP

Table 1  
 Global statistics about electric energy

	Population (million)	Capacity (GW)	Production (TWh)	Electric energy	%	Energy (m.t.o.e)		Electric energy (%)
				kWh per capita		Total primary energy	Electric energy m.t.o.e.	
<b>Global electric energy balance (2005)</b>								
OECD	1020	2000	9000	8800	57	5000	1970	39
East Europe	400	420	1600	4000	10	1070	420	39
DCs	4580	1160	5100	1100	33	3430	1060	31
World	6000	3580	15700	2600	100	9500	3450	37
<hr/>								
		GW	TWh	% of electric energy		Input (m.t.o.e)		%
<b>Fuels for generation of electric energy</b>								
Nuclear		360	2500	16		650		18
Hydro		830	2900	18		250		7
Thermal:		2330	10000	64		2560		72
solids		1100	6100	39		1570		44
oil		430	1350	9		290		8
gas		800	2550	16		700		20
Other		60	300	2		110		3
Total		3580	15700	100		3570		100
<hr/>								
			Nuclear (TWh)	Hydro (TWh)		Thermal and other (TWh)		Total (TWh)
<b>Sources for generation of electric energy (2005)</b>								
OECD			2000	1300		5700		9000
East Europe			240	290		1070		1600
DCs			60	1310		3730		5100
World			2500	2900		10300		15700
<hr/>								
<b>Notes:</b>								
(1) Generation refers to gross generation.								
(2) Capacity refers to net capacity.								
(3) Fuel required for production of electric energy is estimated at 255 g of oil equivalent per kWh for high income and 270g per kWh for East Europe and developing countries.								
(4) Fuel equivalent of hydro generation is equal to the energy content of the generation of electric energy.								
(5) m.t.o.e. = million tons of oil equivalent.								
(6) Terawatt-hour (TWh) = 1000000000 kWh, Gigawatt (GW) = 1000000 kW.								

This calls for a more thorough evaluation of the financial and economic impact of projects in the electricity supply industry. There is a need to move from deterministic assumptions, for studying and modelling of future plans, into risk analysis and mitigation and valuation of uncertainties. Least-cost planning based on deterministic assumptions, doesn't explicitly evaluate economic costs and risks of failure to achieve economic objectives. Therefore, the results of the investment are suboptimal. Also, deterministic methods don't help in understanding the mechanisms that greatly influence investment decisions, like the discount rate and impact of environmental factors.

#### IV. TRADERS OF ELECTRIC ENERGY

Because of price volatility in the bulk supply markets the importance for traders of electric energy has increased during recent years. Correspondingly risk management and protecting against such volatility is one of the critical services offered by traders of electric energy. This is likely to be extended soon from the bulk supply market to the retail competitive markets. Trades of electric energy have access to valuable information (prices, options, etc.), which will also assist their clients in reducing costs, increasing profits and improving performance. They can also split and undertake transaction management and assist in providing liquidity by helping their customer to change position quickly. The services related provided with electric energy by power traders include physical delivery of power or a financial obligation or promise to do so. Such services extend beyond electricity into other energy-related services like natural gas. The delivery may be firm or non-firm, short or long term, one time or stretching over time. Prices may be firm or indexed to other commodities or derived from combination of underlying commodity prices, hence 'derived markets'. Many operators in the business of electric energy won't be able to function successfully in the liberalised competitive for market of electric energy without the help of electric energy marketer. Marketing of electric energy has to investigate customer needs, decide on the prices to offer and the considerable risks and how to

manage them through mitigation and hedging strategies. Small, uncorrelated risks may simply be managed through risk pooling. Aggregating for large count of small risks, which aren't correlated, that is the easiest way to manage risk. As from the beginning of the 21st century, trading of electric energy is growing at a phenomenal rate. In the US the annual volume of wholesale electric energy as commodity trading was anticipated to reach  $2 \times 10^{12}$  \$ by the year 2003 and online transactions to reach  $3.5 \times 10^{12}$  \$ by the year 2005. Such amounts will make electric energy the single largest traded commodity in the US. By 2003 the  $2 \times 10^{12}$  \$ was expected to be about 10 times the retail value of electric energy in trade in the US. This is made possible by the developments in financial trading instruments such as futures, forwards, swaps, and options that allow the same electric energy to be bought and sold several times before they are actually generated and consumed.

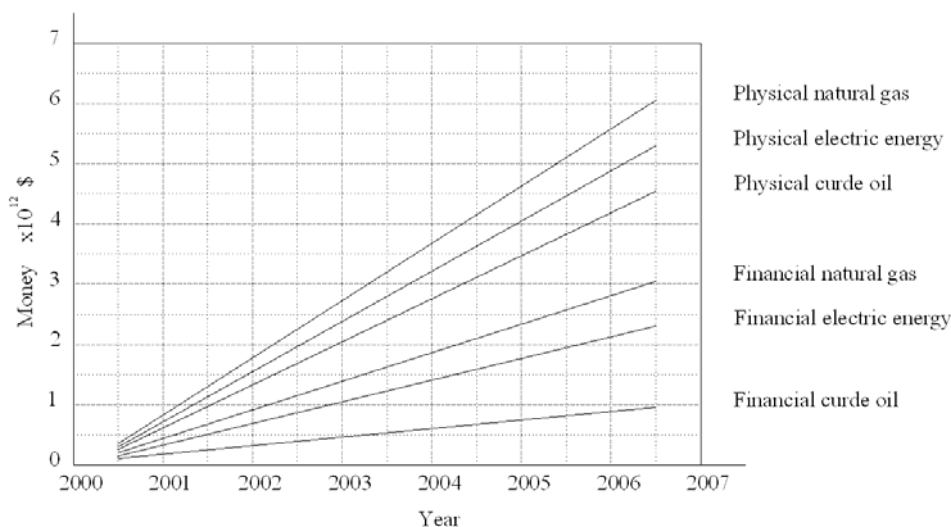


Fig. 3 Online energy trading projections

The volume of trade is typically many times the physical number of kWh generated or consumed because many times electric energy trader will be buying from or selling to another trader. Correspondingly the same kWh is usually bought and sold a number of times until it is finally delivered and consumed. The most important is the quantification, management and hedging of risk.

## V. CONCLUSION

Although the introduction of competition in markets of electric energy is high up the European agenda, there is still some way to go before market integration is complete. As a result, exchanges need to adopt country-by-country strategies, recognising the deregulation agenda and conforming to the competitive environment of each individual country. In essence, exchanges need to adopt a portfolio approach that encompasses corporate strategy, the market environment of each country and the competitive landscape. The legislative environment and regulatory regime of each country need to be considered, as well as the differences in the country's market potential, market participants and competitors. Potential partnerships and alliances can be vital and these opportunities should not be overlooked. The surge in the development and growth of electric energy exchanges around the world has largely been driven by the triumvirate of market forces: legislation, competition and technology. This trend is likely to continue in the short term as legislation in Europe and North America increases the number of new markets opening. In the medium and long term, however, customer demand and increasing competition will be the driving forces. The world's markets of electric energy are fragmented and immature. And the European industry of electric energy, for all its successes, is typical of the global situation. While there are already many exchanges, both of the regulated exchange and the alternative trading systems and varieties, there will be continued growth in the number of exchanges in the short term. Much of this growth in numbers could be reversed in five years' time, as a strong period of consolidation results in the birth of mega-exchanges.

During this time, there will be unprecedented access to new markets through deregulation, better integration of the fragmented markets in Europe and North America, and truly mature markets will develop in these regions, both in electric energy and in gas. This represents significant change and unprecedented opportunity for energy traders and exchanges alike. Being successful requires the right strategy, the right partnerships and the ability to execute those plans globally and on a portfolio basis.

#### ACKNOWLEDGMENT

This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences under the contract No. VEGA 1/0388/13.

#### REFERENCES

- [1] JARET, P.: 'Electricity for Increasing energy efficiency', EPRI J., April 1992, 17, (3)
- [2] GERHOLM, T. R.: 'Electricity in Sweden – Forecast to the year 2050' (Vattenfal, Sweden, 1991)
- [3] KHATIB, H.: 'Economics of reliability in electrical power systems' (Technicopy, Glos, 1978)
- [4] NAKICENOVIC, GRUBLER, and MCDONALD: 'Global Energy Perspectives' (Cambridge University Press, 1998)
- [5] 'Financing worldwide electric powers, can capital markets do the job?'. Resource Dynamic Corporation for US Department of Energy, April 19, 1996
- [6] 'Senior Expert Symposium on Electricity and the Environment', International Atomic Energy Agency (IAEA), May 1991, Helsinki, Finland
- [7] MUNASINGHE, M.: 'Energy Analysis and Policy' (Butterworth, London, 1990)
- [8] BACON, R.W. and BESANT-JONES, J.: 'Global Electric Power Reform, Privatisation and Liberalisation of the Electric Power Industry in Developing Countries', Annual Reviews – Energy and the Environment. The World Bank, 2001
- [9] 'Siemens: Pictures of the Future – Emission Certificate Trading', Siemens Picture of the Future, Spring 2002
- [10] MINCHENER, A.: 'Coal comes clean', IEE Rev. Nov./Dec. 1991, 37 11
- [11] SIMON, M.: 'The road to environmentally compatible power and heat generation', ABB Rev., 3, (91)
- [12] 'Greenhouse gas emissions – the energy dimension', OECD/IEA, Paris, 1991
- [13] BLARK and VEATCH: 'An Attractive Option', Electricity International, May 2002, pp 26–28
- [14] NADEL, S. and GELLER, H.: 'Utility DSM', Energy Policy, April 1996, 24, (4)
- [15] 'Energy, Electricity and Nuclear Power Estimates for the period up to 2020', IAEA, Vienna, July 2001 edition
- [16] 'World Electric Power Plants' (Utility Data Institute, Division of the McGraw Hill, 2nd edn., and later editions, 1996)
- [17] KHATIB, H. and MUNASINGHE, M.: 'Electricity, the environment and sustainable world development', World Energy Council 15th congress, Madrid, September 1992
- [18] BOURDAIRE, J.-M.: 'The link between GNP and energy consumption', Oxford Energy Forum, May 1995, 21
- [19] '1996 World Energy Outlook', IEA, Paris, 2000
- [20] KHATIB, H.: 'Electricity in the global energy scene', IEE Proc. A, June 1993, 140, (1)
- [21] ARGIRI, M. and BIROL F.: 'Renewable Energy', Oxford Energy Forum, Issue 49, May 2002, pp 3–5
- [22] 'World Development Report 2002', The World Bank, Washington DC, 2002
- [23] MULLER-JENTSCH, D.: 'The Development of Electricity Markets in the Euro - Mediterranean Area', World Bank Technical Paper No 491, Washington DC, 2001
- [24] 'Energy and the Challenge of Sustainability – World Energy Assessment', UNDP, UNDESA, WEC, New York, September 2000

# Design of Secure Server using the Micro Kernel and verification

<sup>1</sup>Peter GONDA, <sup>2</sup>František JAKAB, <sup>3</sup>Ivan KLIMEK, <sup>4</sup>Jaroslav LÁMER

<sup>1,2,3,4</sup> Department of Informatics, Faculty of Electrical Engineering and Informatics,  
Technical University of Košice, Slovak Republic

<sup>1</sup>peter.gonda@student.tuke.sk, <sup>2</sup>frantisek.jakab@tuke.sk, <sup>3</sup>ivan.klimek@tuke.sk,  
<sup>4</sup>jaroslav.lamer@tuke.sk

**Abstract** — This paper deals with secure server model design using formal verified microkernel OKL4. Secure server design is based on detailed analysis of leading microkernels from L4 microkernel's family. Approach to model composition into components and their continuous integration is described. Usage of paravirtualized operating system Wombat, based on microkernel in cooperation with designed secure server, is also depicted. This paper additionally includes benchmark.

**Keywords** — Secure server, microkernel, OKL4, Fiasco, IPC

## I. INTRODUCTION

In information technologies is generally known that new kinds of software are frequently invented. Can we thus assume that given the increasing number of different software products, their security remains at the same level? Bearing this in mind it is now important to deal with this issue – mainly since we live in an information society, whose trends include global digitalization and automatization of previously manual tasks. This provides easier access to information however on the other hand it means easier misuse of this information. Thus, security of information should be an elementary part of all processes.

Underestimating risks and neglecting security in this field can lead to fatal consequences, such as identity or property theft, etc. While the development of technologies is moving rapidly forward, it is open to many security risks. In spite of this, it is still being implemented in everyday life in the form of phone software, ATM software, etc. Security often takes a back seat and possible threats are underestimated.

To be able to achieve the highest possible level of security, it is necessary to build on solid foundations, such as the lowest level of the operating system - the kernel. It is the cornerstone of both the management of physical resources of a computing system and the management of allocation of computing time to individual processes.

## II. GOALS

The aim of the paper is to design a model of a secure server capable of providing adequate protection from attacks frequently occurring in a monolithic kernel. For this, a verified microkernel is a proper solution. Using microkernel, security issues were taken into consideration in its very design and various mechanisms to minimize any possible risks we utilized.

In dealing with this issue, a formally verified microkernel OKL4 will be utilized, with paravirtualized Wombat operating system running on top of it. In parallel with Wombat, a secure server as a microkernel thread is to be running this will allow communication with the secure server by means of a single available mechanism in the microkernel - IPC messaging.

Messaging between microkernel threads will be carried out exclusively by the microkernel. In order for the threads to communicate, they need to own a microkernel object on an IPC channel, which is the only mean of communication. The microkernel thus guarantees the impossibility of communication between a certain thread and a counterfeit thread, etc.

### III. ANALYSIS

Kernel is a computer program managing the input/output requirements of programs and converts them into instructions processing the data for the processor (CPU) or other electronic devices of the computer. The kernel is a crucial part of an operating system [1]. When a computer program (a process, in this case) performs kernel requests, these are called system calls. Individual kernel formats differ in their ways of handling system calls (time-sharing) and resources. For example, a monolithic kernel of an operating system performs all system instructions in the same address space in order to improve system performance [6]. Kernel represents the lowest level of software abstraction on hardware. Its role is to manage system resources and integrate applications into data processing on the hardware level. In a monolithic kernel, drivers, file systems and many other functions are parts of the kernel of an operating system. These services require a privileged access to system resources and physical memory and only the kernel has direct access to these resources [1] [9].

The microkernel manages the majority of the operating system in the background of the user space, in order for the operating system to be more modular and easier to maintain [2]. For software programmers, a kernel interface is a low-level abstraction. Microkernel is a micronucleus of an operating system (OS), which only performs the basic functions of the OS, such as memory management, process (thread) management and inter-process communication (IPC). Its name signifies its own basic characteristic – it is not included into the family of robust kernels with hundreds of thousands of source code lines but rather runs over smaller number which increases readability. This feature also helps make troubleshooting in the microkernel much easier. Kernel is a crucial part of the OS. In general, an OS consists of 2 parts:

- kernel interface (privileged mode),
- user interface (non-privileged mode).

Without this division, it would be impossible to ensure security between individual processes. Kernels can be divided into two groups - monolithic kernel and microkernel.

#### A. Monolithic Kernel

This type of kernel is characterized by its executable code running in the same memory space, in a so-called kernel space. It provides services such as inter-process communication management, file systems, device management, process management, interruption interception, input-output communication, drivers etc. Since all these services are provided in kernel interface, i.e. privileged mode, incorrect implementation of a particular element can lead to security violation. For example, if a security violation occurs in a driver, it could threaten the security of the OS because of the shared address space, where an attacker could gain access to system information and change it this would endanger the entire OS performance. Allocation of basic services into kernel interface results in these disadvantages:

- kernel size increases,
- scalability decreases,
- sustainability decreases,
- troubleshooting effort increases.

#### B. Microkernel

The concept of the microkernel depicted on Fig 1 was supposed to reduce the kernel to inter-process communication, process management and memory management.

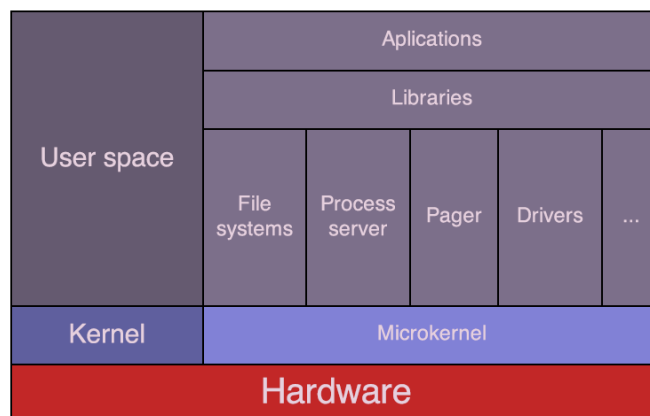


Fig. 1 Structure of a microkernel OS



The rest of the OS, such as file system, drivers, etc. are implemented as so-called servers which are executed in user interface and use their own address spaces, which ensures a high level of security. These servers can communicate with each other solely by means of microkernel IPC, which also contributes to higher security [7]. No shared memory is implemented, as it could be used to gain unauthorized access to data which is being used by another process and could be misused.

#### IV. SOLUTION AND RESULTS

Security and reliability of computer systems is closely connected with security and reliability of the kernel used in the operating system. OKL4 is a third-generation microkernel. In the third generation of microkernels, optimization improved IPC, process management and memory management, which significantly increased performance. This microkernel is the only formally verified microkernel [5]. While designing a secure server, security and isolation of execution of the main components into their own address spaces should be crucial, in order to avoid leaking or misuse of sensitive information [4]. Its formal verification guarantees that there are no programming errors. OKL4 was used in this thesis due to its characteristics [3].

In the application design itself, existing methods and types of microkernels were taken into consideration. The key point of the design is the isolation of individual components into separate address spaces in order to prevent unwanted direct access into the address space of the other process. In other words, it is necessary to run an application independently as another OKL4 thread, which is represented as a TCB microkernel object. This object carries key information about the thread, such as collection of *options* owned by the thread, information about available memory for a particular thread, thread priority (used for planning and opening processor time) etc.

The primary component of the design using OKL4 is Wombat Linux. *Wombat* is a paravirtualized Linux OS, running on top of OKL4 microkernel [3]. It utilizes the *iwana* memory library for management of the entire memory available to it. It also has a modified *libc* library, which uses *iwana* for creating new Linux processes actually implemented as OKL4 threads but running in the same address space as the Wombat Linux kernel.

The main reason for its use is simple hardware management, thanks to implemented necessary drivers, such as network card drivers and file systems. It also contains implemented TCP/IP protocol used in network communication which could be easily used in the future.

Secure server is an application running in its own address space, parallel to Wombat running in another address space, which has access to peripherals and has implemented drivers. This will enable its communication with said devices. These two components can only communicate via IPC communication but only indirectly via microkernel calls, which mediate the actual information exchange.

The communication is based on the fact that for two threads to be able to communicate, they need an address space or a structure created before running, in which received data will be stored in order for the thread to be able to access it. It is called an IPC buffer and it is located in the TCB object of a given thread. During IPC communication, microkernel copies these structures between communicating threads. By means of microkernel API it is possible to read from and write into the IPC buffer.

##### A. Microkernel Component

For the application to be easily scalable in the future, it is necessary to divide related elements into smaller parts. In this case, it is useful to divide related elements into so-called components. A component is an OKL4 application running in its own address space separated from the rest of the components.

This application is designed as an independent component for paravirtualized Wombat Linux. A secure server is a component implemented as a dependency during Wombat operating system compilation. During the compilation itself, all components are compiled independently, i.e. the output of the component is a one *elf* file, which is later used in linking and creating a bootable image of the OS. Communication between individual processes is possible only via IPC calls. To connect the components, a so-called kernel object *endpoint* is used. If an OKL4 process wants to communicate with another OKL4 process, it needs to have a reference to a particular endpoint. A single endpoint can be referenced by 3 or more processes [8]. OKL4 object allocation is performed by *iwana* library, which is also used by *wombat-supervisor* component. It is a component utilized for initialization of various kernel objects necessary for running Wombat Linux. It creates abstract IPC channels between processes, *capabilities* for hardware

and individual channels of *time* server (implemented as a component as well).

*Wombat-supervisor* creates kernel objects on the basically defined components implemented as structures which are located in the *system\\_one\\_wombath.h* header file in a directory with header files of this component. To ensure the creation of an IPC communication, a *SYSTEM\\_CONNECTIONS* variable is defined as a field containing *connection\\_t* structures, which includes information about both sides of the IPC communication channel, abstract channel type (*SYNC\\_ENDPOINT*, *VIRTUAL\\_IRQ* etc.) and writing and reading rights from a given communication channel.

To boot the component, it was necessary to create a function which allocates a given component, sets its name and assigns appropriate devices and OKL4 thread priority to it. Then this function returns the newly created component as a *component* structure.

### B. Linux Component

To enable communication between Wombat and microkernel component, it is necessary to implement a component running in the same address space as the Wombat Linux, in order for it to be able to use its functions, such as time measurement, dynamic memory allocation, system load measurement and so on.

Booting a Linux component is performed during Linux kernel initialization. In order to ensure proper functioning and use of Linux services, it is crucial to make sure a component runs after the Linux kernel is fully started. It is necessary to wait until the kernel starts IRQ interception, sets system time, etc. A Linux component is booted directly from the *post\\_init()* initialization function, which is located in the *init/main.c* Linux source code file before the initialization of user mode.

### C. Comparison of monolithic kernel and microkernel - time of IPC message

Time it took to send a IPC message between Linux monolithic kernel and microkernel was evaluated.

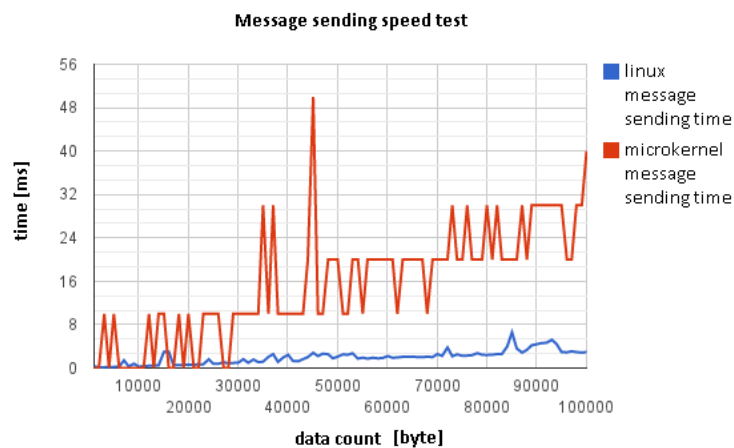


Fig. 2 First message sending speed test between monolithic Linux kernel and microkernel

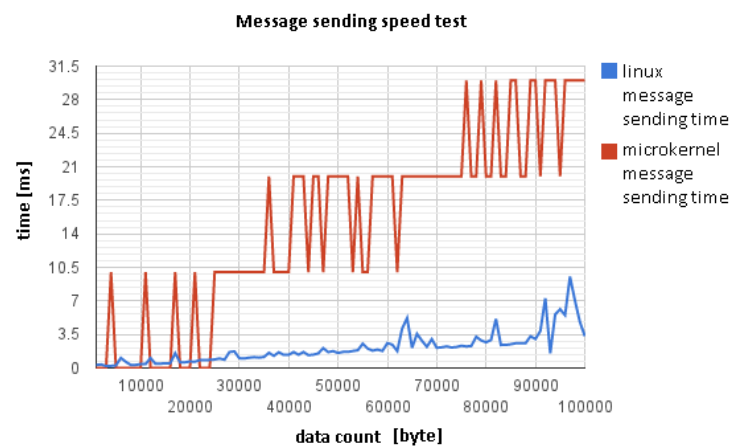


Fig. 3 Second message sending speed test between monolithic Linux kernel and microkernel

All measured results are illustrated in Fig. 2. The communication time, needed by microkernel (in red color) is much higher than monolithic kernel. The second measurement is made with the same configuration and conditions and is evaluated in Fig. 3.

#### V. CONCLUSION

This paper presents a secure server model designed with the first formally verified OKL4 microkernel and paravirtualized Wombat operating system in order to separate individual applications into different address spaces. This prevents unwanted interference of one application into the part of the memory belonging to another application, aiming to misuse processed data.

In the design, the *OKL4* microkernel was used. The main reason was that it is currently the only formally verified microkernel from the most modern L4 microkernel family. It ensures the absence of programming errors, which could pose a significant threat to the security of the entire system and it also shows considerable increase of speed of IPC messaging, as opposed to other microkernels from the L4 family. Other members of the family include for example Pistachio and Fiasco, which are implemented in C++ language, in which the formal verification process is more complicated than in *OKL4*, since that is implemented in C language.

The performed measurements proved that IPC messaging in a monolithic kernel is faster than IPC messaging in a microkernel. However, in terms of security, a microkernel solution is more useful, even at the cost of decreased speed, as opposed to monolithic kernels.

Paper is aimed to bring an experimentally verified design of a secure server which formally enables the creation of a secure application run by the designed system. Further improvement could be achieved by using a newer Linux kernel in Wombat, as well as separate implementation of microkernel component which would be responsible for network communication and would provide network services for other microkernel components. Alternatively, an extended *cpio* archive could be created, enhanced with applications and services for remote file system connection such as NFS, packaging system for simpler application installation, and more.

#### ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] Daniel P. Bovet and Marco Cesati. *Understanding the Linux Kernel*. O'Reilly, Sebastopol, California, October 2000.
- [2] Harvey M. Deitel. *An Introduction to Operating Systems (2Nd Ed.)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990.
- [3] Lukas Hanel. Systematic setup of compartmentalized systems on the seL4 microkernel. Master's thesis, Technische Universität Dresden, February 2010.
- [4] Gernot Heiser, T. Murray, and G. Klein. It's Time for Trustworthy Systems. *Security Privacy, IEEE*, 10(2):67-70, 2012.
- [5] Gerwin Klein, June Andronick, Kevin Elphinstone, Gernot Heiser, David Cock, Philip Derrin, Dhammika Elkaduwe, Kai Engelhardt, Rafal Kolanski, Michael Norrish, Thomas Sewell, Harvey Tuch, and Simon Winwood. seL4: formal verification of an operating-system kernel. *Commun. ACM*, 53(6):107-115, 2010.
- [6] Theodore A. Linden. Operating System Structures to Support Security and Reliable Software. *ACM Comput. Surv.*, 8(4):409-445, December 1976.
- [7] Benjamin Roch. Monolithic kernel vs. Microkernel, 2011.
- [8] J.S. Shapiro. Vulnerabilities in synchronous IPC designs. In *Security and Privacy, 2003. Proceedings. 2003 Symposium on*, pages 251-262, 2003.
- [9] Karim Yaghmour. *Building embedded Linux systems - concepts, techniques, tricks, and traps*. O'Reilly, 2003.

# Bezdrôtový prenosový systém s viacerými vstupmi a viacerými výstupmi

Iveta GLADIŠOVÁ

Katedra elektroniky a multimediálnych telekomunikácií  
Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, SR

iveta.gladisova@tuke.sk

**Abstrakt** — Článok sa zaoberá bezdrôtovým prenosovým systémom MIMO využívajúcim princíp diverzity, ktorý je v dnešnej dobe už súčasťou niekoľkých medzinárodných štandardov. Poskytuje pohľad na MIMO kanál ako taký, jeho kapacitu v porovnaní s tradičnými komunikačnými kanálmi, a popisuje rôzne typy diverzít.

**Kľúčové slová** — MIMO, prenosový kanál, diverzité domény, kapacita kanála.

## I. ÚVOD

V dôsledku zvyšujúcich sa nárokov na kvalitu služieb a prenosových rýchlostí telekomunikačných systémov boli vyvinuté systémy s viacerými vstupmi a viacerými výstupmi t.j. MIMO (Multiple Input/Multiple Output). Princíp MIMO systému spočíva v zvyšovaní prenosových rýchlostí a zabezpečení vyššej spoľahlivosti, ako tomu bolo v prípade systémov s jedným vstupom a jedným výstupom. V čase, keď boli MIMO systémy opísané v strede deväťdesiatych rokov Gerardom Foschinim a inými [1], prekvapujúca efektívnosť šírky pásma takýchto technológií sa zdala byť v rozpore so Shannonovým limitom. Žiaden rozpor však v skutočnosti neexistoval, pretože diverzita a spracovanie signálu použité v MIMO transformujú jednoduchý kanál typu bod-bod do viacerých paralelných alebo maticových kanálov, čo spôsobuje násobenie kapacity. MIMO ponúka vyššie prenosové rýchlosti a spektrálnu efektívnosť. Táto výhoda je tak jasná, že mnoho štandardov už inkorporovalo MIMO technológiu. ITU používa MIMO vo vysokorýchlostnom paketovom prístupe (HSPDA - High Speed Downlink Packet Access), časti UMTS štandardu. MIMO je taktiež časťou štandardu 802.11n, používanom bezdrôtovými smerovačmi, a taktiež 802.16 pre mobilný WiMax, ktorý je využívaný mobilnými sieťami. LTE štandard takisto inkorporuje MIMO [11].

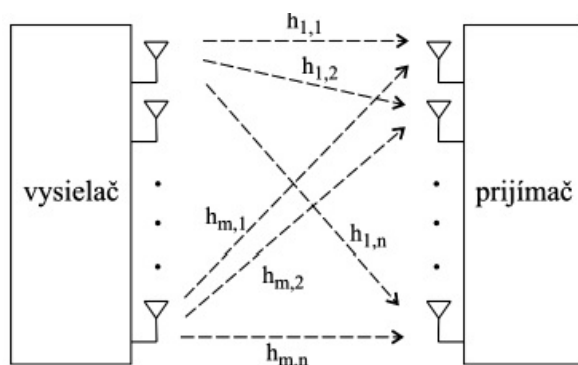
## II. MODEL MIMO SYSTÉMU

Predstavme si komunikačný systém, v ktorom je v každom časovom intervale  $t$  vysielaný signál  $s_{t,m}$ , pre  $m=1, 2, \dots, m_T$  z  $m_T$  vysielacích antén. Všetky  $s_{t,m}$  vyslané signály prechádzajú bezdrôtovým komunikačným kanálom, na konci ktorého sú prijímané  $n_R$  prijímacími anténami. Bezdrôtový komunikačný systém s  $n_R$  prijímacími anténami, kde každý výstup z kanála je lineárnou superpozíciou prenášaných signálov, predstavuje vstupy kanála rušené šumom. Všetky vytvorené páry z vysielacích a prijímacích antén predstavujú jedinú cestu signálu od vysielateľa k prijímaču. Zisk cesty medzi  $m_T$  vysielacou a  $n_R$  prijímacou anténou je vyjadrený koeficientom  $h_{m,n}$ . [4]. Na obr. 1 je diskretný model MIMO kanála.

Prijatý signál  $r_{t,m}$  v čase  $t$  získaný prijímacou anténou  $n_R$  môžeme na základe tohto modelu zapísať nasledovne:

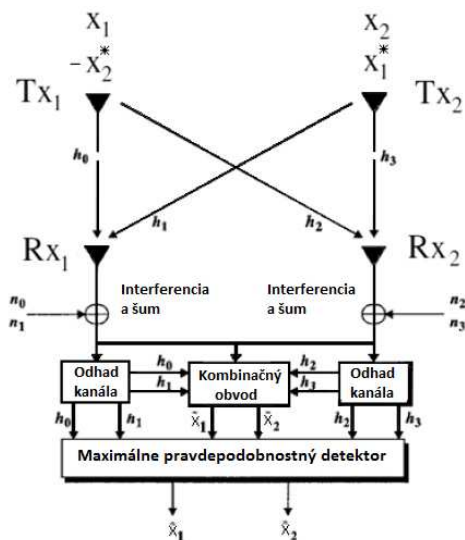
$$r_{t,m} = \sum_{n=1}^{n_R} h_{m,n} \cdot s_{t,m} + \eta_{t,n} \quad (1)$$

kde  $\eta_{t,n}$  predstavuje šumovú vzorku prijímacej antény  $n_R$  v čase  $t$ . Táto rovnica reprezentuje kópie signálov zo všetkých vysielacích antén, ktoré sú sčítane na každej prijímacej anténne. Aj keď sú verzie utlmených signálov v prijímacích anténach zmiešané, dosiahneme pomocou tohto systému diverzity zisk v dôsledku prítomnosti viacerých kópií vyslaného signálu na prijímači [2].



Obr.1 Diskrétny model MIMO kanála.

Na obr.2 je znázornený model MIMO systému s dvomi vysielacími a dvomi prijímacími anténami. Vysielame rovnaké symboly z každej antény na rovnakej frekvencii, ktoré sú prijímané na dvoch prijímačoch. Každá prenosová cesta od vysieláča k prijímaču má určitú stratu/zisk a daný kanál môžeme touto stratou charakterizovať. (Prenosová cesta môže v skutočnosti pozostávať z veľa komponentov, ale je charakterizovaná len začiatočnými a koncovými bodmi). Keďže všetky štyri kanály v tomto prípade nesú rovnaký symbol, zabezpečuje to diverzitu tým, že silnejšie kanály „nahrádzajú“ slabšie. V Tab.1 je uvedená kódovacia a vysielacia sekvencia, v Tab.2 je definícia kanála medzi vysielacími a prijímacími anténami a v Tab.3 sú prijaté signály na dvoch prijímacích anténach [3].



Obr. 2 Model MIMO systému s 2 vysielacími a 2 prijímacími anténami [3].

Tab. 1  
 Kódovacia a vysielacia sekvencia

	$T_{X1}$	$T_{X2}$
čas $t$	$X_1$	$X_2$
čas $t+T$	$-X_2^*$	$X_1^*$

Tab. 2  
 Definícia kanála medzi vysielacími a prijímacími anténami

	$R_{X1}$	$R_{X2}$
$T_{X1}$	$h_0$	$h_2$
$T_{X2}$	$h_1$	$h_3$

Tab. 3  
 Prijímané signály na dvoch prijímacích anténach

	$R_{X1}$	$R_{X2}$
čas $t$	$r_0$	$r_2$
čas $t+T$	$r_1$	$r_3$

Pre prijaté signály platí:

$$\begin{aligned} r_0 &= h_0 x_1 + h_1 x_2 + \eta_0 \\ r_1 &= -h_0 x_2^* + h_1 x_1^* + \eta_1 \\ r_2 &= h_2 x_1 + h_3 x_2 + \eta_2 \\ r_3 &= -h_2 x_2^* + h_3 x_1^* + \eta_3 \end{aligned} \quad , \text{ kde} \quad (2)$$

$\eta_0, \eta_1, \eta_2$  a  $\eta_3$  sú komplexné náhodné premenné reprezentujúce šum a interferencie na prijímači a  $h_0, h_1, h_2$  a  $h_3$  sú impulzné odpovede kanála, niekedy nazývané kanálové koeficienty. Kombinačný obvod na Obr.2 vytvára dvojicu signálov:

$$\begin{aligned} \overline{x_1} &= h_0^* r_0 + h_1 r_1^* + h_2^* r_2 + h_3 r_3^* \\ \overline{x_2} &= h_1^* r_0 - h_0 r_1^* + h_3^* r_2 - h_2 r_3^* \end{aligned} \quad , \quad (3)$$

ktoré sú následne privedené do maximálne pravdepodobnostného detektoru, v ktorom sa zisťuje, aký symbol bol s najväčšou pravdepodobnosťou vyslaný [3].

### III. KAPACITA SISO A MIMO KANÁLOV

Tradičná komunikačná linka, ktorú nazývame kanál s jednoduchým vstupom a jednoduchým výstupom, tzv. SISO kanál (Single-In-Single-Out), má jeden vysielateľ a jeden prijímač. Ale namiesto jedného vysielateľa a jedného prijímača ich môžeme použiť niekoľko, t.j. MIMO kanál. Kapacitu SISO linky, označenú ako  $C$ , ktorá je špecifikovaná počtom bitov, ktoré môžu byť prenesené cez kanál a sú merané metrikou (b/s/Hz), môžeme vyjadriť ako [4],[5]:

$$C = \log_2(1 + SNR) \quad (4)$$

Kapacita SISO linky je iba funkciou SNR (z angl. Signal-to-Noise Ratio) kanála a tento kapacitný vzťah bol stanovený Claudom Shannonom [6] a je taktiež nazývaný informačno-teoretická kapacita. SNR v tejto rovnici je definované ako celkový výkon vydelený výkonom šumu. Ak je SNR je 10 dB, kapacita kanála je 3,64 b/s/Hz. Desať násobné zvýšenie výkonu, teda SNR s 20 dB, zvýši kapacitu na 6,65 b/s/Hz, čo je menej ako dvojnásobok pôvodnej kapacity. Sto násobné zvýšenie výkonu zvýši kapacitu kanála iba na 9,96 b/s/Hz, čo je približne trojnásobok pôvodnej kapacity. Kapacita sa zvyšuje ako logaritmická funkcia SNR, čo je pomalé zvyšovanie. Zjavne zvyšovanie kapacity akýmkoľvek výrazným faktorom si žiada enormné množstvo výkonu v SISO kanáli [11].

MIMO nám to umožňuje zvyšovať kapacitu lineárne s funkciou výkonu (desaťkrát zvýšime výkon, desaťkrát zvýšime kapacitu). Kapacita MIMO systému sa zvyšuje lineárne s počtom antén, pričom SISO/SIMO/MISO systémy sa zvyšujú iba logaritmicky [3]. Prenosová kapacita MIMO kanála závisí od korelácie signálov medzi jednotlivými anténami, ktorá môže prispieť k nárastu ako aj k zníženiu kapacity prenosového kanála. Kapacitu kanála vyjadrieme pomocou Shannonovej informačnej teórie. Kapacitu kanála môžeme vyjadriť ako maximálnu prenosovú rýchlosť dát k ľubovoľne malej pravdepodobnosti chyby. Pridaním  $m_T$  vysielacích a  $n_R$  prijímacích antén docielime rozšírenie diverzity na oboch stranách kanála a tak dostaneme MIMO systém s kapacitou, ktorú môžeme potom vyjadriť ako [5], [7]:

$$C_{MIMO} = \log_2 [\det(I_{n_R} + (SNR/m_R) H^H H)] \quad (5)$$

kde  $I_{n_R}$  vyjadruje jednotkovú maticu pre prijímaciu anténu,  $H$  je kanálová matica, v lit. nazývaná „Channel Information“, v ktorej sú zapísané impulzné odpovede kanála  $h_{m,n}$  a  $(\ )^H$  vyjadruje transponovanú konjugáciu. Ak je počet vysielacích antén  $m_T$  veľký, potom kapacitu MIMO systému môžeme vyjadriť aj ako:

$$C_{MIMO} = n_R \log_2 (1 + SNR) \quad (6)$$

Použitím priestorového multiplexovania získame vyššiu kapacitu alebo väčšiu šírku pásma použitím viacerých antén na oboch stranách prenosového kanála pri zachovaní energetickej náročnosti. Vďaka diverzitému zisku dosiahneme lepšiu spoľahlivosť spojenia pri rovnakých prenosových rýchlostiach nezávisle od únikov [9]. Výhody MIMO systému sú

- vyššia kapacita (bit/s/Hz)
- lepšia kvalita prenosu
- lepší odhad polôh užívateľa
- lepšie pokrytie .

Motivácie pre zavedenie MIMO systémov sú

- dopyt po bezdrôtových komunikáciách s vysokou dátovou rýchlosťou (blížiacou sa k 1 Gb/s)

- návrh bezdrôtových liniek s vysokou prenosovou rýchlosťou pri zachovaní dobrých QoS vlastností.

#### IV. ZÁKLADNÉ METÓDY MIMO SYSTÉMOV

Pre efektívnejšie využitie MIMO systémov, ktoré disponujú viacerými prijímacími a vysielačmi anténami v telekomunikačných systémoch boli vyvinuté dve základne metódy:

- Priestorový multiplexing ( SM- Spatial Multiplexing)
- Princíp diverzity

##### ***Priestorový multiplexing***

Táto metóda spôsobuje zvýšenie prenosovej rýchlosti pri zachovaní rovnakej šírky pásma, v ktorom je signál vysielať s rovnakým vysielačím výkonom, pričom je dosiahnuté zvýšenie spektrálnej účinnosti a teda lepšie využitie kapacity MIMO systémov. Cieľom tejto metódy je prenos maximálneho množstva informačných dát. Spomenutá metóda je vhodná pre prostredie s priamou viditeľnosťou, kde sa vyskytuje malý počet únikov, rozmery zariadenia a celkový počet antén nie je obmedzený. Základným princípom priestorového multiplexovania je rozdelenie postupnosti informačných bitov do množiny niekoľkých subpostupností, ktoré sú spracovávané samostatne. Použitím priestorového multiplexovania je  $N$  nezávislých symbolov vysielať z  $N$  antén, čo dovoľuje bližšie priblíženie k maximálnej prenosovej kapacite MIMO systémov.

##### ***Princíp diverzity***

Priestorová diverzita vylepšuje spoľahlivosť komunikačných liniek, ktoré môžu byť dosiahnuté vysielačmi a prijímačmi viacerých kópií tej istej informácie pomocou rôznych komunikačných kanálov s únikmi. Používa prídavnú priestorovú diverzitu na obmedzenie strát signálu v prenosovom kanáli s únikmi. Posielať kópií signálov po rôznych komunikačných cestách sa zníži možnosť znehodnotenia signálu pôsobením nízkej pravdepodobnosti výskytu únikov vo viacerých komunikačných kanáloch súčasne. Počet prijatých nezávislých kópií informácie nazývame *diverzitným ziskom* systému. Vysielačie a prijímačie antény vytvárajú diverzitný zisk, ktorý je rovný súčinu vysielačích a prijímačích antén v MIMO systémoch, ktoré vytvárajú  $N$ - $M$  nezávislých komunikačných ciest. Priestorovú diverzitu je vhodné použiť v prostredí, kde je vysoký počet únikov na vylepšenie spoľahlivosti SISO systémov, pričom prenosová rýchlosť sa nemení a počet antén v MIMO systéme je využívaný len na zlepšenie spoľahlivosti. Systémy využívajúce princíp diverzity nachádzajú uplatnenie v komunikačných systémoch s malým počtom vysielačích antén, ktoré vysielať signál s pomerne malým odstupom od šumu (SNR). Podstatnou nevýhodou a podmienkou týchto systémov je, že vysielať signály z jednotlivých antén musia byť dostatočne dekorelované, antény dostatočne vzdialené medzi sebou, pričom vzdialenosť by mala predstavovať minimálne polovicu vlnovej dĺžky prenášaného signálu. V MIMO systémoch, ktoré sú zamerané na zvýšenie vysielačieho a prijímačieho diverzity, sa kanálové kódovanie označuje podľa toho, v ktorej oblasti sa diverzita aplikuje.

##### *Časová diverzita*

V tomto prípade sa opätovné vysielať signálu vykonáva v čase za pomoci kanálového kódovania a prekladacích metód. Kľúčom k úspechu u tohto typu diverzity je, že kanál musí byť dostatočne premenný v čase. Túto diverzitu je možné použiť v prípade tam, kde je doba koherencie kanála malá v porovnaní s dobou trvania prekladaného symbolu. V tomto prípade je zaistené, že prekladaný symbol bude nezávislý od predchádzajúceho symbolu. Tento spôsob nám taktiež poskytuje úplnú repliku originálneho symbolu.

##### *Frekvenčná diverzita*

Tento typ diverzity poskytuje kópiu originálneho signálu vo frekvenčnej doméne. Môžeme ho použiť iba v prípade, ak koherenčná šírka pásma kanála je menšia v porovnaní so šírkou pásma signálu.

##### *Priestorová diverzita*

Niekedy nazývaná aj ako anténna diverzita. Kópia signálu je v tomto prípade vysielať viacerými vysielačmi anténami k prijímaču. Vzdialenosť jednotlivých vysielačích antén sa volí tak, aby sa zamedzilo koherentným únikom. Priestorovú diverzitu môžeme ďalej rozdeliť do dvoch podskupín, podľa toho kde je diverzita aplikovaná, či už na strane prijímača alebo vysielača:

- *Prijímač – priestorová diverzita*

Maximálny kombinovaný podiel (Maximum ratio combining) je často používaný diverzitný systém v prijímačoch pre vylepšenie kvality signálu. Pri použití v mobilných telefónoch je veľmi drahý a ťažko implementovateľný. To je hlavným dôvodom prečo sa využíva skôr vysielačie diverzita.

- *Vysielacia – priestorová diverzita*

Je ľahšie implementovateľna, stačí ju implementovať v základňovej stanici. Pri použití tejto diverzity zavedieme do signálu určitú redundanciu, ktorú ide ľahko kontrolovať. Obecne táto technika vyžaduje kompletne znalosti prenosového kanála pre vysieláč. Tieto požiadavky môžeme odstrániť použitím priestorových kódovacích schém ako napr. Alamoutiho kódovanie [4], [8].

Obecne efektívnosť každého diverzitetného systému závisí od toho, aby prijímač zaistil nezávislosť vzoriek základného signálu, ktoré sú vysielané. Diverzitetný systém musí optimálne zlúčiť rôznorodé signály tak, aby vybral signál s najlepšou kvalitou [9].

## V. ZÁVER

Témou článku bolo vo všeobecnosti popísať využitie princípu diverzity v prenosových systémoch. Technológia MIMO využíva priestorovú diverzitu. Bol popísaný model a charakteristika MIMO komunikačného kanála s dvomi vysielacími a prijímacími anténami. Popísaná bola kapacita MIMO kanála v porovnaní so SISO kanálom. Dá sa konštatovať, že kapacita MIMO systému sa zvyšuje lineárne s počtom antén, pričom pri SISO systémoch ide o logaritmické zvyšovanie. MIMO teda ponúka vyššie prenosové rýchlosti a spektrálnu efektívnosť.

## LITERATÚRA

- [1] Foschini, G. - Ganss, M.: *On limits of wireless communications in a fading environment when using multiple antennas*. In: Wireless Personal Communications, vol. 6, 1998, pp. 311-335.
- [2] Žalud, V.: *Moderní radioelektronika*, 2004, ISBN 80-86056-47-3.
- [3] Alamouti, S. M.: *A Simple transmit diversity technique for wireless communications*. IEEE Journal on Select Areas in Communications, Oct. 1998. Dostupné na internete: <<http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=730453&url=http%3A%2F%2Fiee>>
- [4] Jankiraman, M.: *Space-Time Codes and MIMO systems*. London: Artech House, 2004. Dostupné na internete: <<http://read.pudn.com/downloads166/ebook/765643/Space-Time%20Codes%20and%20MIMO%20Systems.pdf>>
- [5] Gesbert, D. - Akhtar, J.: *Breaking the barriers of Shannon's capacity: An Overview of MIMO wireless systems*. Department of Informatics, University of Oslo, 2002.
- [6] Shannon, C. E.: *A Mathematical Theory of Communications*. In: Bell Systems Technical Journal, vol. 27, 1948, pp. 379-423 and 623-656.
- [7] Telatar, e.: *Capacity of Multiantenna Gaussian Channels*. In: European Transactions on Telecommunications, vol. 10, no. 6 (1999), pp. 585-595.
- [8] Alamouti, S. M.: *A Simple transmit diversity technique for wireless communications*. IEEE Journal on Select Areas in Communications, Oct. 1998. Dostupné na internete: <<http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=730453&url=http%3A%2F%2Fiee>>
- [9] Harish G.: *Multiple-Input-Multiple-Output (MIMO) Systems*, Dostupné na internete: <[www.cse.buffalo.edu/~qiao/cse620/fall04/mimo.ppt](http://www.cse.buffalo.edu/~qiao/cse620/fall04/mimo.ppt)>
- [10] Langton, Ch. - SKLAR, B.: *Finding MIMO*. Dostupné na internete: <<http://www.complextoreal.com/chapters/mimo.pdf>>
- [11] Gladišová, I.- Virág, L.: *Prenosové systémy využívajúce princíp diverzity*. In: Posterus.sk. Roč. 6, č. 3, s. 1-7, 2013.



# Improving Reliability of Arena Platform for Automated Assessments

<sup>1</sup>*Miroslav Biňas*

<sup>1</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>miroslav.binas@tuke.sk

**Abstract** – This paper reports on the design and usage of the *Arena* platform at the *Technical university of Košice*. Services involved in the evaluation process have been used by the students during the summer semester of 2014. Paper describes an issue and its solution related with the reliability of evaluation service, when the results with feedback are returned to the students immediately.

**Keywords** – automatic assessment, reliability, service oriented architecture

## I. INTRODUCTION

The need for automated assessment tools comes from the large numbers of students enrolling in programming courses at many institutions [1]. The number of students usually raises in the first semesters of study. In such situation is management of the class and assignments evaluation unmanageable. The motivation to use tools providing automated assessment become very straightforward.

This was also the main reason at *Faculty of Electrical Engineering and Informatics Technical University of Košice* in the introductory programming course in the second semester of study. The average number of students is around 450 and to stay objective with every single assignment is very hard. Automated assessment tools can help to solve this situation very easily.

The first attempt started in the year 2012, when we tried some automatization in the process of assignments submission with our local installation of LMS *Moodle*. Then next year we tried to include to the submission process version controlling system *Git* and created some unit tests for automated evaluation. Finally this year, we created system based on the *service oriented architecture* with the name *Arena*, which helped us in the process of submission and evaluation of students assignments.

The issue, which rises, was during the final exam. Students were programming their solutions, and the tests were running on demand - they could evaluate their solution whenever before the end of exam. If the number of students was above 40, system became overloaded and didn't evaluate some solutions until some system resources used during evaluation were available again.

This paper presents solution for increasing reliability and provides scalability of *Arena's* evaluation service.

## II. RELATED WORK

Many tools providing some level of automated assessment have been already developed. There exist several reviews of available tools, such as [2], [1] or one of the latest [3]. One of the main reasons, why there are still new tools developed, authors of mentioned reviews consider fact, the creators are trying to solve usually local problem or their need for some specific course [1]. And because of their local "orientation", it is very hard to spread them and use in other courses or institutions. Another reason for constat development of new tools is the lack of open-source systems providing automated assessment [3].

Some of the tools are more known than others, such as *BOSS* [4], *CourseMarker* [5] or *Web-CAT* [6], because they were used in more than in just one institution or course. But the main problem still remains - they are suitable for specific type of tests, problems or courses.

The architecture is in most cases *client-server* and in some cases are clients provided as desktop applications. This is for example the situation of previously mentioned systems *BOSS* [4] and *CourseMarker* [5]. The disadvantage of such solution is the need to keep updated versions of local clients.

Different architecture than *client-server* was discussed in the paper [7]. This time authors discussed *service oriented architecture* for the e-assessments. We have also used *service oriented architecture* (SOA) for development of *Arena platform*.

### III. ARENA PLATFORM

*Arena* platform is based on *SOA* architecture and all of its components are self-contained and independent services. Closer description and experiences from usage of the platform during the summer semester of the year 2014 is described in paper submitted for publication [8]. Currently there are three services involved in the platform:

- 1) **Structure Checker** - The purpose of this service is to check the structure of submitted student's archive. Input for the service contains the student's archive itself and its description in the form of *XML* file.
- 2) **Arena** - This is the main service of the whole platform. It is responsible for evaluation of student's assignments. Input for the service is student's solution in the form of source codes and the test description in *JSON* format.
- 3) **Caesar** - This service is for presentation purposes only. The result of every test returned from the *Arena* service is presented in the form of *HTML* web page through this service.

We used these services in several use-cases. All of them are described in more details in paper submitted for publication [8]. One of the use-cases, which we used for final exams is illustrated on figure 1.



Fig. 1 Block diagram of final exam

In this situation we found the bottleneck of the evaluation service. The service itself works as expected, but the problem raises with the number of requests for evaluation. Specially in the introductory courses, where the number of students with some basic programming experiences is high, the probability for presence of infinite loops in their code is very high. If evaluation process consists of list of for example 10 tests with maximum duration of 2s for each of them, in the worst case the duration of whole test (and the whole evaluation) is 20s. As a result, service can act as overloaded due to the number of requests it is currently processing. The service is provided through *Apache* web server and every single request is represented as an open connection between the client and service. Finally, the number of opened connections is limited in the web server configuration. This situation can be annoying for the students taking the test, because with the high number of students taking the test, the probability to get no results from evaluation service raises.

### IV. IMPROVING RELIABILITY

To avoid the problem with reliability of evaluation service, the service architecture must be changed. Reliability itself can be increased by rising the number of simultaneously running evaluation services on different hosts. With the increased number of evaluation services, the problem with overloading can be decreased. And with the option to add new or to remove an old evaluation service from running system, the scalability feature can be achieved.

In such case, the communication with the service can't be direct as it is in the original architecture. The communication between the client and the evaluation service must be separated through another layer. As a solution for such situation is to create some sort of *task queue*, where the tasks (student's solutions) will be stored for later processing [9].

There exist plenty of implementations today for some sort of queues. The most known and common today are for example *Apache ActiveMQ* [10], *RabbitMQ* [11], or *Redis* [12].

One of the required feature among the existing implementations for task queues, is the implementation of *publish/subscribe* pattern [13], which is originally known as *observer* design pattern [13]. The concept of *publish/subscribe* is characterized by *listeners* subscribing to *channels*, with *publishers* sending binary string messages to channels. Anyone listening to a given channel will receive all messages sent to that channel while they're connected and listening [9].

In the case of evaluation service, the *publisher* will be client sending the student's code for evaluation. This request will contain the source code or metadata describing its location and it will be stored to the specific *task queue* and the information about new task will be spread to the listeners over specific *channel*. Client will get *task ID* in return, so it will be able to retrieve the result based on this *ID* later.

As a *listeners* subscribed to this specific *channel* will be services provided evaluation. First service, which receives message about new task in *task queue* will mark this task as *being evaluated* and starts with it's evaluation. When the evaluation is complete, it stores the results and marks the task as *evaluated*. Now, when the client asks for the result based on the *task ID*, it will receive the result of the evaluation. After the result is retrieved by the client, it is removed from the *task queue*.

Improved use-case presented on figure 1 used during the final exam is illustrated on the figure 2. This time with pattern *publish/subscribe* included.

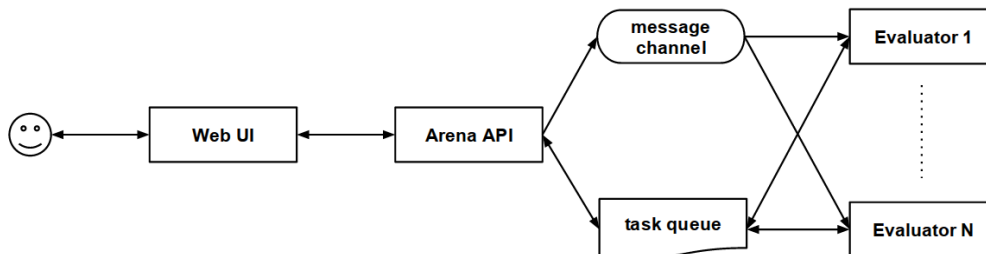


Fig. 2 Publish/subscribe pattern integrated into *Arena* service

Original *Arena* service was separated into two parts:

- 1) *Arena API Service*, which provides the original API for evaluation of student's solutions, but is extended for method of getting the resource of final evaluation.
- 2) *Evaluator* service or process, which is responsible for the evaluation itself.

This two parts are interconnected with the *message channel*, which is used for spreading the information about the changes in *task list*. *Arena API Service* acts as a *publisher* and *Evaluators* act as a *subscribers*.

## V. TESTING

Students were using *Arena* platform during whole semester. There were almost 500 students involved in the semester and they had to solve 4 assignments. According to the evaluation process, there were two types of evaluations:

- 1) *On-Demand Evaluation* - In this case, student's solution was evaluated immediately after it was submitted for evaluation.
- 2) *Scheduled Evaluation* - Students were submitting their solutions into VCS and in defined periods of time (e.g. every three hours) the evaluation process was executed.

There was only one assignment of type one and three assignments of type two. During semester there were no issues related to the reliability of the *Arena service*. Some issues were expected during the period of solving the assignment of type one, but it was easy enough to raise any issue. Together, there was over 25000 evaluations during the semester.

As was mentioned earlier, the issues with reliability raised during final exam, when there was more than 40 students on the same term simultaneously - some of the evaluations didn't complete because of the high number of requests for evaluation. Student's had to try to evaluate their solution several times to get some.

Fixing this issue we started after the last exam where the issue was identified. For *message channel* and *task queue* we have used *Redis*, which is popular persistent in memory key-value database [12]. The tests were handled on some sort of test data with two *listeners* only, but with the ability to scale the service by adding new and removing existing *evaluators*. If, for example, one *evaluator* is capable to handle maximum of 20 clients and their requests simultaneously, with adding *N* new *evaluators* is the whole system able to handle  $N * 20$  requests from clients simultaneously.

## VI. CONCLUSION

Automated testing and assessment can be very important tool for teacher, because it brings time efficiency and objectiveness in the evaluation process. Many tools at many institutions have been developed to reach this two goals. One solution based on the *service oriented architecture* was also built at *Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice*.

Based on our experiences, the *Arena* platform helped us a lot and students were also confronted with their results right after the evaluation. During the semester and introductory programming

course they also learned some basics of *test driven development*, because they were fixing their issues based on the results of automated tests.

Together we have over 34000 records of student's submissions and evaluations. That means, that average number of submissions per student is around 70. And to handle all of 70 submissions of every student by hand is really unmanageable.

We identified important reliability issue, when the number of simultaneously evaluated students is too high. Description of this issue with the changes made in the *Arena* service, were described in this article. With this change, the *Arena* service became more reliable and scalable.

#### REFERENCES

- [1] A. Pears, S. Seidman, L. Malmi, L. Mannila, E. Adams, J. Bennedsen, M. Devlin, and J. Paterson, "A Survey of Literature on the Teaching of Introductory Programming," 2010.
- [2] K. Ala-Mutka, "A survey of automated assessment approaches for programming assignments," *Computer Science Education*, pp. 83–102, 2005.
- [3] P. Ihantola, T. Ahoniemi, V. Karavirta, and O. Seppälä, "Review of recent systems for automatic assessment of programming assignments," *Proceedings of the 10th Koli Calling International Conference on Computing Education Research - Koli Calling '10*, pp. 86–93, 2010. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=1930464.1930480>
- [4] M. Joy, N. Griffiths, and R. Boyatt, "The boss online submission and assessment system," *Journal on Educational Resources in Computing*, vol. 5, no. 3, pp. 2–es, Sep. 2005. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=1163405.1163407>
- [5] C. A. Higgins, G. Gray, P. Symeonidis, and A. Tsintsifas, "Automated Assessment and Experiences of Teaching Programming," vol. 5, no. 3, pp. 1–21, 2006.
- [6] M. A. P.-Q. S.H.Edwards, "Web-cat: automatically grading programming assignments," *ITICSE '08: Proceedings of the 13th annual Conf. on Innovation and technology in computer science education*, pp. 328–328, 2008.
- [7] D. R. M. Amelung, P. Forbrig, "Towards generic and flexible web services for e-assessment," in *ITICSE '06: Proceedings of the 11th annual SIGCSE Conf. on Innovation and technology in computer science education*. ACM, 2008, pp. 88–92.
- [8] M. Biñas, "Web services in the process of evaluation of student's projects," in *Booktitle*, 2014.
- [9] J. L. Carlson, *Redis in Action*. Manning Publications, 2013.
- [10] "activemq.apache.org," <http://activemq.apache.org/>, (Visited on 07/17/2014).
- [11] "Rabbitmq - messaging that just works," <http://www.rabbitmq.com/>, (Visited on 07/17/2014).
- [12] "Redis," <http://redis.io/>, (Visited on 07/17/2014).
- [13] R. J. J. V. Erich Gamma, Richard Helm, *Design patterns: elements of reusable object-oriented software*. Reading, Mass. : Addison-Wesley, 1995.

# Data collection and evaluation for the purpose of assistive applications for road transport

<sup>1</sup>Jana Šťastná, <sup>2</sup>Miroslav Biňas

<sup>1,2</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>jana.stastna@student.tuke.sk, <sup>2</sup>miroslav.binas@tuke.sk

**Abstract** – In recent years, trends in the use of mobile technologies for various purposes are rapidly increasing. One of them is provision of assistive applications for road transport. These applications are often based on additional information about traffic conditions. This work is dealing with collection and evaluation of data about traffic signs. It provides an insight into problems of collecting data through community volunteers using crowdsourcing. The main contribution of this work is a possible solution for evaluating and updating of data which are obtained by method mentioned above.

**Keywords** – Crowdsourcing, assistive applications, mobile technology, traffic signs, data evaluation

## I. INTRODUCTION

Increase in availability of mobile technology has opened up possibilities for prevention of traffic accidents. Drivers may become more aware of current traffic conditions. In this regard a system using actual traffic signs data would be useful, however, it requires large database of proper information that would be regularly updated. The problem is obtaining and maintaining of such database.

The subject of this work is to point out possibilities of obtaining a large amount of additional information with an inexpensive and fast way. This paper presents one of them - a method called crowdsourcing - that appears to be a promising solution of the problem.

The second problem is processing of the large amount of received information. It is necessary to expect some errors and inaccuracies when collecting data from a large quantity of volunteers. This work proposes also the evaluation procedure that can provide gradual verification and updating of information and exclusion of erroneous data.

The goal of this work is to design a procedure for evaluation of additional information about traffic signs, that may be obtained from a large quantity of voluntary contributors. Proposed method can be applied in assistive applications for road transport.

## II. RELATED WORK

Various commercial companies but also voluntary organizations are focused on collecting additional information also known as points (or places) of interest. A majority of commercial companies requires payments for the use of data they provide. On the other hand, organizations such as *OpenStreetMaps* provide useful information free of charge for further use. It is a difficult task that can be fulfilled by involving voluntary contributors. The way of engaging large numbers of people in this activity is known as *crowdsourcing*.

The concept of *crowdsourcing* was mentioned for the first time in 2006 in an article *The Rise of Crowdsourcing* [1] written by *Jeff Howe*. It was described as a “new organizational form, in which various resources or activities are transmitted via public challenges from within the organization to another outside party - heterogeneous internet community” [2] [3].

Basic components of crowdsourcing are defined in [2] as follows:

- 1) organization which requires completion of a specified task,
- 2) community that is willing to participate voluntarily in solving the task,
- 3) on-line environment that will allow operations necessary to fulfill the task and interaction between the organization and volunteers,

4) mutual benefit for both groups involved - the organization and volunteers.

From the point of view of data collection, this method has especially following advantages:

- 1) lower costs compared to hiring specialized employees,
- 2) high number of participants,
- 3) a competition may arise between participants, which will encourage them to greater activity,
- 4) there is a possibility to get independent data and feedback from users of different locations,
- 5) it is suitable for obtaining quantitative results.

When trying to deploy this method some disadvantages may also occur:

- 1) crowdsourcing requires interesting propagation from the start, otherwise it may not be successful,
- 2) management of large numbers of volunteers causes complications and requires time,
- 3) it is difficult to ensure cooperation between participants when it is needed,
- 4) system can not rely on participants; they can cancel participation at any time,
- 5) results may be less reliable, compared to work of experts,
- 6) the method is suitable only for simple tasks that do not require special knowledge, skills and experience.

Social networks and mobile technologies are improving conditions for deployment of crowd-sourcing. Mobile application that can detect occurrence of traffic signs using image recognition is suitable for collecting desired information.

Recognition of traffic signs has already been addressed in several articles, which propose quite effective procedures, e.g. [4] and [5]. Some of them focus directly on deployment in intelligent cars [6], which draw much attention in recent years. Number of *Peaks Algorithm* presented in [7] seems to be very interesting. This algorithm is aimed at traffic signs of triangular shape and reaches exceptionally high accuracy - 100 percent. It could be subjected to further examination and with appropriate changes adjusted for traffic signs of other shapes.

### III. SOLUTION AND RESULTS

#### A. Data collection

By combining modern mobile technologies and crowdsourcing can be created a system that will ensure collection and updating of data regarding traffic signs.

Additional information like

- GPS location,
- GPS accuracy, or
- speed,

can be obtained from GPS module in mobile device (smartphone) and together with identifier of recognized sign and timestamp of recognition they will create an unprocessed entry about traffic sign, defined as  $U$ . The entry will be sent from the mobile application to the web service that will evaluate it. All the data that will be evaluated by service are shown in figure 1.

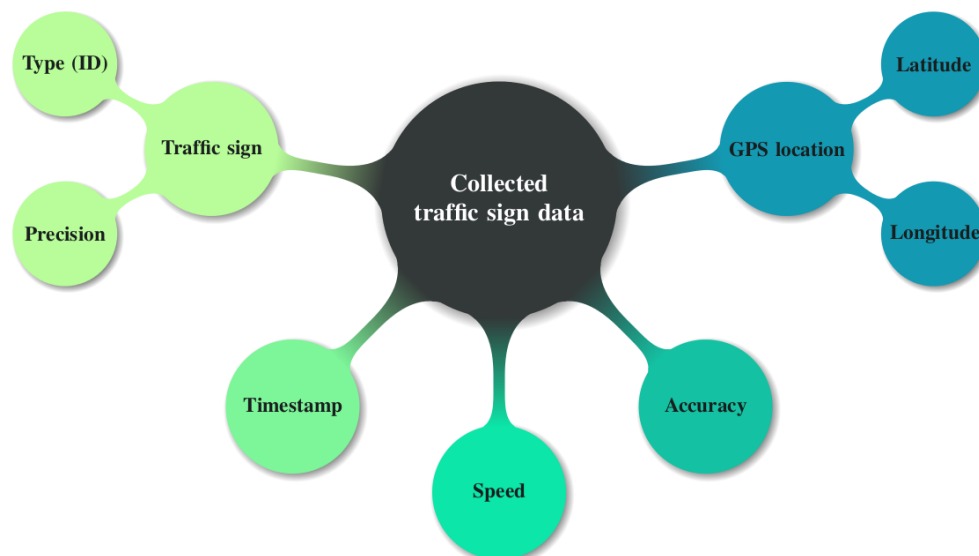


Fig. 1 Traffic sign with additional data that will be evaluated by web service

Data from unprocessed entry  $U$  are defined as follows:

- identifier (type) of traffic sign  $U_{id}$ ,
- precision of image recognition  $U_p$ ,
- timestamp of recognition  $U_t$ ,
- speed at which sign was recognized  $U_s$ ,
- accuracy of GPS location  $U_a$ ,
- location
  - latitude  $U_{lat}$ ,
  - longitude  $U_{lon}$ .

It is important to ensure that the data load when transferring collected entries  $U$  is as low as possible and that the client mobile application and web service use a common communication format.

### B. Data evaluation

There are several ways to accomplish the data evaluation by web service. One of them is proposed in the following section.

It is assumed that when the large group of volunteers will participate in data collection, the service will receive large quantity of data that will be stored in a database. These entries will be later evaluated in order from the oldest to the newest. This order will ensure that the updating procedure corresponds to real changes.

Processed entries, defined as  $P$ , will be stored in a separate database table. They will contain partly different data than unprocessed entries  $U$ . Following information will be essential in entries  $P$ :

- identifier (type) of traffic sign  $P_{id}$ ,
- degree of verification  $P_v$ ,
- location
  - central latitude  $P_{lat_c}$ ,
  - central longitude  $P_{lon_c}$ ,
- area of traffic sign defined by
  - minimal latitude  $P_{lat_{min}}$ ,
  - maximal latitude  $P_{lat_{max}}$ ,
  - minimal longitude  $P_{lon_{min}}$ , and
  - maximal longitude  $P_{lon_{max}}$ .

The evaluation procedure is made by cycle of steps that are as follows:

- 1) From the database that holds unprocessed entries  $U$  is taken the oldest entry.
- 2) Entry  $U$  is compared with processed entries  $P$ . The database is searched for such entry  $P$ , for which the following applies:

$$\begin{aligned}
 P_{id} &= U_{id}, \\
 P_{lat_{min}} &\leq U_{lat}, \\
 P_{lat_{max}} &\geq U_{lat}, \\
 P_{lon_{min}} &\leq U_{lon}, \\
 P_{lon_{max}} &\geq U_{lon}.
 \end{aligned}$$

- 3) If a matching entry  $P$  is found (figure 2, stage 1 on the left, green marker A), it means that the entry  $U$  (red marker 1) belongs to the area of entry  $P$ , which will be then updated.
  - a) According to the value of  $U_p$ , the value for  $P_v$  will increase of 1 point resp. 2 points, in the case that  $U_p = 100$ .
  - b) Boundaries for area of traffic sign from entry  $P$  will update.
    - i) According to values  $U_a$  and  $U_s$ , boundaries will be calculated for area with its center specified by values  $U_{lat}$  and  $U_{lon}$  (figure 2, stage 2 on the right).
    - ii) If the area in entry  $U$  is bigger than area in entry  $P$ , boundaries will not change (figure 3, stage 3 on the left, stage 4 on the right and figure 4, stage 5 on the left).
    - iii) If dimensions of the area are smaller or equal to dimensions of the area in the entry  $P$  (figure 4, stage 6 on the right), new coordinates for  $P_{lat_c}$  and  $P_{lon_c}$  will be calculated as average of values  $U_{lat}$ ,  $U_{lon}$  and  $P_{lat_c}$ ,  $P_{lon_c}$  (figure 5, stage 7 on the left).
    - iv) In that case, boundaries for area of the record  $P$  will update according to the entry  $U$  (figure 5, stage 8 on the right).
- 4) If no entry  $P$  corresponding to the entry  $U$  is found, new entry  $P$  will be created.



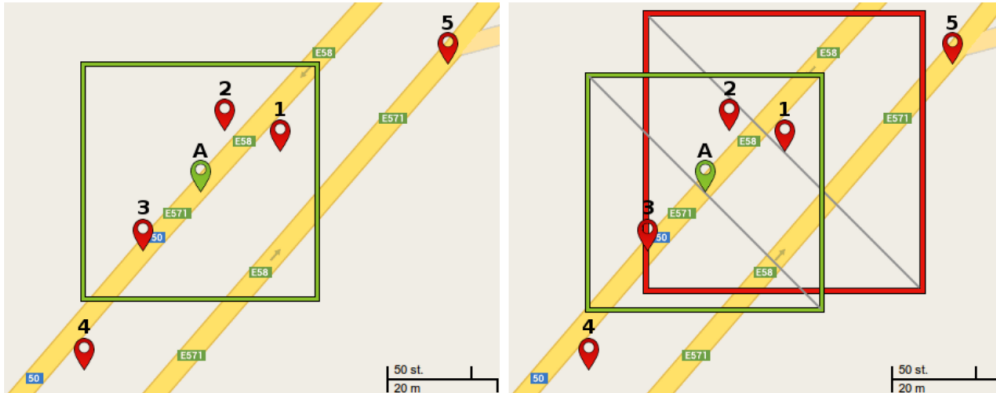


Fig. 2 Data evaluation, stage 1 (left) and stage 2 (right)

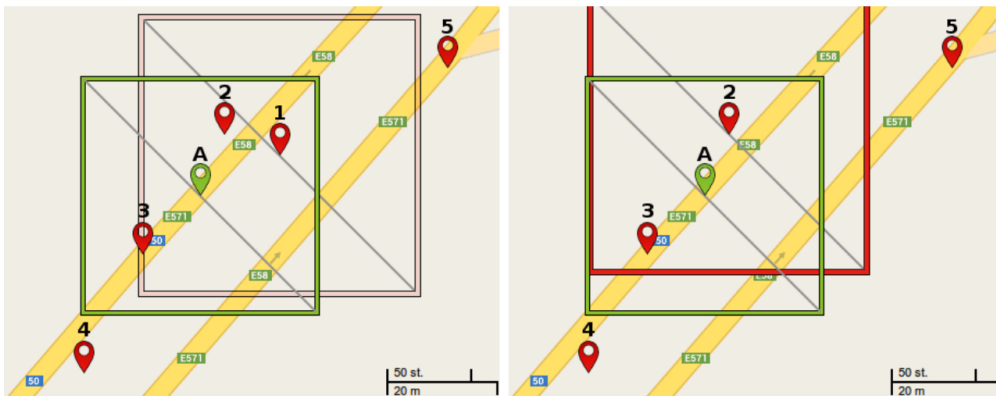


Fig. 3 Data evaluation, stage 3 (left) and stage 4 (right)

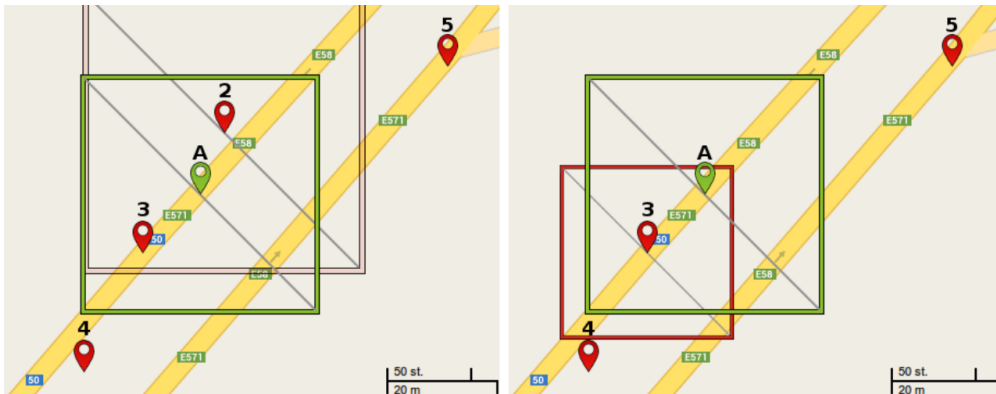


Fig. 4 Data evaluation, stage 4 (left) and stage 6 (right)

- a) Value  $P_{id}$  will be set to  $U_{id}$ .
- b) Degree of verification  $P_v$  will be set to 1.
- c)  $P_{lat_c}$  will be set to value  $U_{lat}$  and  $P_{lon_c}$  to value  $U_{lon}$ .
- d) Then a distance defined as offset will be calculated from values  $U_a$  and  $U_s$ , to create borders for area around location of the detected sign. Values of borders will be calculated as follows:

$$\begin{aligned}
 P_{lat_{min}} &= U_{lat} - offset, \\
 P_{lat_{max}} &= U_{lat} + offset, \\
 P_{lon_{min}} &= U_{lon} - offset, \\
 P_{lon_{max}} &= U_{lon} + offset.
 \end{aligned}$$

- e) Resulting entry  $P$  will be saved to database.

5) The entry  $U$  will be marked as evaluated.

The more entries of given type of traffic sign in given location will be evaluated, the highest



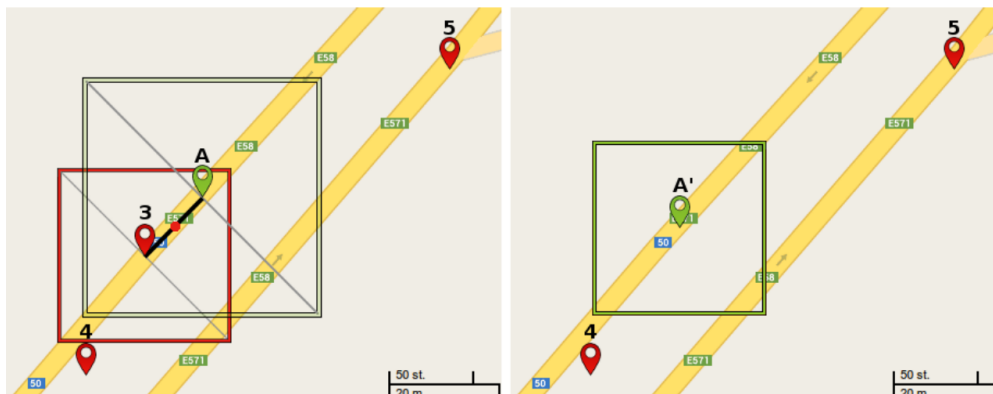


Fig. 5 Data evaluation, stage 7 (left) and stage 8 (right)

will the degree of verification  $P_v$  become.

The next stage of evaluation, the opposite to the above, is decrementation of value  $P_v$ . Entries will be evaluated in ascending order according to their timestamp:

- 1) Entry  $P$  will be inspected for the time of its last update. If the time elapsed since the last update exceeds a given value  $T_{min}$ , the degree of verification  $P_v$  will update.
  - a) If  $P_v \geq 1$ , then  $P_v$  will decrease of 1 point or different given value.
  - b) Else if  $P_v = 0$  and the time elapsed since the last update exceeds a value  $T_{max}$ ,  $T_{max} > T_{min}$ , the entry  $P$  will be erased from the database. Otherwise it will remain stored in the database.

Variables  $T_{min}$  and  $T_{max}$  depend on the size of the group collecting data, frequency of collection and the amount of collected data. In the beginning it would be appropriate to set variables to a high value, so that evaluated data would not be erased from the database too early.

#### IV. CONCLUSION

The goal of this work was to design a procedure, which will provide verification and updating of information and exclusion of potentially erroneous data regarding traffic signs.

The proposed procedure may be useful with the active participation of a large group of users which will gather required data. It is less financially demanding than obtaining data from a specialized commercial company and brings independent and diversified results.

Disadvantages that may emerge are complications and time demands associated with managing a large group of voluntary contributors. To attract a crowd it requires attractive propagation or reward system for participation. Moreover, the reliability of data obtained in this way is not ensured. It is therefore more suitable for projects with low risk. However, if tasks for contributors are simple and do not require any special knowledge, skills or abilities, the collection and evaluation of additional information with this proposed method may be successful.

#### ACKNOWLEDGMENT

This work is the result of the project implementation: *Development of the Center of Information and Communication Technologies for Knowledge Systems* (ITMS project code: 26220120030) supported by the *Research & Development Operational Program* funded by the *ERDF*.

#### REFERENCES

- [1] J. Howe, "The rise of crowdsourcing," *Wired*, vol. 14, June 2006. [Online]. Available: [http://archive.wired.com/wired/archive/14.06/crowds\\_pr.html](http://archive.wired.com/wired/archive/14.06/crowds_pr.html)
- [2] D. C. Brabham, *Crowdsourcing*. MIT Press, 2013.
- [3] E. H. C. Michael S. Bernstein, "Leveraging online populations for crowdsourcing," *IEEE Internet Computing*, vol. 16, no. 15, pp. 10–12, 2012.
- [4] Z. Z. Yangping Wang, Jianwu Dang, "Traffic signs detection and recognition by improved rbfn," in *International Conference on Computational Intelligence and Security*. IEEE Computer Society, 2007, pp. 433–437.
- [5] M. Avila, "Detection and recognition of traffic signs using gabor filters," in *34th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE Computer Society, 2011, pp. 554–558.
- [6] B. K. Zhixie Chen, Jing Yang, "A robust traffic sign recognition system for intelligent vehicles," in *Sixth International Conference on Image and Graphics (ICIG)*. IEEE Computer Society, 2011, pp. 975–981.
- [7] A. Hamandi, "Recognition of triangular traffic signs using the number of peaks algorithm," in *2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*. IEEE Computer Society, 2012, pp. 210–214.

# Basics of formal specification of component programming

*Valerie NOVITZKÁ, William STEINGARTNER, Martina BENČKOVÁ,  
Pavol MACKO*

Department of Computers and Informatics, Faculty of Electrical Engineering and  
Informatics Technical University of Košice, Slovak Republic

valerie.novitzka@tuke.sk, william.steingartner@tuke.sk, martina.benckova@tuke.sk,  
pavol.macko@tuke.sk

**Abstract** – Nowadays, there are still more and more approaches and processes how to make faster and easier development with lower cost. Because of this there grows up a count of components, software elements, that we can use again in different program systems. However each component or library is needed to compose with the rest of system different way. It depends on standard that component is using. On the other hand there is a problem to understand what component exactly means. People are wrong or confused in terminology. Because of this we tried to unify the terminology. We think this can help to find an unified formal description what can lead to simplification of creating a formal models. We offer our view to this problems.

**Keywords** – Component- oriented programming, interactions, composition

## I. INTRODUCTION

In last years, new approach of software development, component-oriented programming, became popular. By few words it is composed program system from existing software or hardware components. At the first sight component-oriented programming can be understand as a higher level of object-oriented programming, but there are many important differences. Component-oriented programming enables reusing of prepared components in quite different program systems. Reusability requires some independence of components and some rules how to compose and deploy them to achieve a working program system.

Customized systems used to be developed from zero but this approach is high cost and needs long time. Also customer wants some changes many times during implementation and , only with help of some tools and libraries. The procedure libraries can be considered as the first example of the simple form of component programming. This approach has some advantages, e.g. simple modifications, employing knowledge and skills of clients, etc. But software developed in this manner has high cost and it cannot be optimally used out of given institution. Involving some new functions in such systems, e.g. web access, local or global interoperability can be very difficult for programming teams. Present world is characteristic with rapidly changing requirements and this classical programming approach is not able to keep up with it.

Now, there are several so called standard components in many component libraries and they can be used in different systems as needed. Such purchased software can be adapted to user requirements only slightly, it often enforces great changes in the work mode in user institution. But all administration, evolution and interoperability is out of users. Among such known companies providing components of wide spectrum belong OMG (Object Management Group), CORBA (Common Object Request Broker Architecture), Microsoft COM (Component Object Model) and CLR (Common Language Runtime) and many others. These companies compete on market; they have no common rules and no common standards for provided components. Development of various components, their accessibility and reusability leads to the situation when component-oriented programming becomes a necessity, the main programming method in software engineering.

There are many publications and web pages about component programming, composition of components, and methods for solving arisen problems. Most of them deal with implementation details and issues, only very small number of them tries to provide some principal, exact formal approach. Also in the terminology, as in every rapidly evolved discipline, is some confusion.

Therefore our effort is oriented to formulate unified terminology, to define particular basic notions and their properties, and to define formal description of basic principles in component-oriented programming. This paper tries to analyze basic notions and characteristics, and to formulate our ideas and considerations to formal description and modeling component-oriented program systems.

## II. WHAT COMPONENT MEANS

The concept of component has many different definitions. First of all, we have to consider differences between objects and components, later we mention some other analogies and at the end of this section we formulate our definition of component.

Many people do not distinguish components in component-oriented programming from objects and/ or classes in object-oriented programming. Both need some access points to be possible to provide their services. These access points are some types. To be possible communication between objects and components there are some interactions. These interactions can be described by some patterns and frames. The basic difference between components and objects we mention below.

The main difference between component oriented programming and object oriented programming is that component oriented programming places importance on interfaces and composition while object oriented programming on classes and objects [13][15][27].

An object can be viewed as

- a unit of instantiation having unique identity sufficient to its identification;
- it has externally observable state.

A component

- is an independent deployable entity;
- has to have all important features, it must be deployed whole, not partially;
- implementation details are hidden from user;
- encapsulates its behavior, only interface is visible, not implementation details;
- has involved interactions with environment by well-defined typed ports;
- can be a subject of separate compilation;
- has no externally observable state.

Nowadays, there are many definitions of component. For instance, among newer formulations in [3][4][23] belong the following characterization: a component is a software implementation that can be deployed independently and it is a subject to composition by third parties. This definition, and any other, does not express all important properties of components. Sometimes, a component is compared with black box, because of hiding implementation details, it can be reused and it can have some rules to be deployed in program system.

Components can be considered as some classes known from object oriented paradigm and their deploying and instantiations lead to some kind of objects. However, components have more properties than objects, even a component can involved several objects [13][28].

Often, to guarantee functionality of component in given program system, some own user code, additional information and external files have to be used. Components can be generic, i.e. substituting parameters by appropriate arguments enable their using for different purposes [2][9][24].

The world of components is still quite chaotic. We try to characterize some so called stable common features as some starting points for formal approaches. We use the following terminology:

- a visible part of a component we call *interface*;
- every component has one or more access point (they can serve for input or output of data).  
There are several names like interface, connector, port, placeholder, etc. We prefer the name *port* which provide best expressing of its actual meaning;
- every port enables to send or receive data of some *type*.

Every component has to have some additional information for its successful deployment in program systems, namely:

- contracts;
- context dependencies.

Components are hardware of software elements that can be added into some program system by composition. Composition is defined on the base of interactions between components. To be this composition successful, contracts should be stated to formulate some necessary conditions. Obviously, every component has several ports of different types for sending or receiving data from or to other components.

*Contracts* create a common basis for successful composition and interaction between components in any program system. In the present, some standards for interfaces are preparing for

these purposes. It can be compared with the situation before the standards of open systems was formulated. The basic conditions stated by contracts involve:

- correlation between ports constructing data and the ports extracting these data;
- requirements to ports of other components that should be satisfied for making given component working;

Contracts can be considered as a specifications of non-functional requirements. Their role is to state obligations for achieving desirable behavior of components in program systems [?], [?]. The first formal specification of contracts follows from algebraic specifications of abstract data types, i.e. signatures and axioms formulated in some logical system extended with some constraints specific for given component. Among simple examples of contracts we can list component interoperability, pre- and post conditions of Hoare's logic, invariants, etc. [8][14][18][31].

Every component should specify own requirements on environment in which it has to work. These context dependencies involve:

- composition context dependencies - requirements on environment for successful composition;
- deployment context dependencies - possible platforms (hardware and software) where component can work.

Due to different application domains there are many ways how to build program systems from components. Most of the components can be considered as data structures with explicit or implicit data types, subprograms, collections of subprograms, etc. One component can require some modification in another one component. This relation between new and existing components is one of form of dependencies. A simple example of dependencies between components is in [29]. Let  $A$  and  $B$  be components with corresponding explicitly typed data structures. Authors classify known and frequently occurred composition context dependencies to the following kinds:

- *data dependencies*: a value of  $A$  can influence data of  $B$ , or data of  $A$  can be used for computing  $B$ ;
- *type dependencies*: type definitions and their changes in  $A$  can influence data types in  $B$ ;
- *subprogram dependencies*: executing some procedure or function of component  $A$  by calling with data from  $B$  causes this kind of dependencies;
- *source file dependencies*: if some source file serves as a common source for both components  $A$  and  $B$  in program system;
- *source location relationships*,
- *time and space dependencies*, and many others.

Deployment context dependencies express hardware and software platform for successful deploying of components. Here belong hardware architectures, operating systems and their versions, representations of built-in data, etc. Deployment context dependencies become important in implementation process and we think they are hardly formalizable.

After this short analyzing and classifying of contracts and dependencies we can extend the definition of components with the following part: A software component is a unit of composition with contractually specified access points and explicit context dependencies. Contracts and composition context dependencies are stable parts in component-based programming and they seem to be the first adepts for rigorous formalization.

### III. COMPONENT COMPOSITION

Composition of components requires constructing composition model. It can be considered as a basic architecture expressing the roles of components their interconnecting, relations between them and the rules of composing components together. Relations between components are called *interactions* [6].

It is not at all clear what a composition model for program system constructed from components should be [25]. For any complex system a model is needed. It creates a basis for independence and cooperation of components on some abstract level. It can be characterized as a holistic view of a program system, defining invariants, i.e. properties characterizing all the actual program systems constructed on the base of this model. A model can be based on principal considerations on functionality and it should state some decision policy ensuring correct interoperability of components. In such complex and large systems as component-based program systems are, it is obvious to construct models in hierarchical mode with several layers.

Certainly, in any model we abstract from a computer architecture or implementation details. For our purposes we consider a model as a collection of

- component frameworks and
- the rules of interoperability between these frameworks.

Component frameworks contain specification of interfaces and typed ports together with decision policy on the component level. The rules of interoperability state the conditions for successful cooperation between components in the form of defining interactions.

In the following subsection we try to classify known and frequently occurred interactions in the frame of component model.

#### A. Interactions

A component is often characterized as a black box with only visible part called interface. All information, i.e. data structures, data types, procedures and ports without their implementation details constitute an interface [30]. Only the content of an interface defines what can be used in program system after deploying a component. An interface is often specified as an abstract data type. Ports are very important part of any interface because cooperation between components can be performed only through corresponding ports. Ports serve as end points of interactions, they enable transfer of data of some type in required direction [20].

It is interesting that different publication provide different classification of interactions. In the following literature [29][25][16][1] authors form various patterns of interactions. On this base we try to classify interactions by three criteria:

- interactions based on pre- and post- conditions;
- interactions that are dealing with concurrency and nondeterminism;
- interactions based on what makes choice and what follows-up.

The first group encapsulates interactions specified on the base of Hoare's pre- and post-conditions. If  $c_1$ ,  $c_2$  and  $c_3$  are preconditions and postconditions, respectively,  $p$  is input or output port, we can distinguish the following interaction patterns:

- *send pattern sequence*: before executing function *send* it is satisfied pre-condition  $c_1$ , after it post-condition  $c_2$  is satisfied. This pattern is illustrated in Figure 1;
- *pre- blocking send*: in this pattern a message is sent through port  $p$ . After pre-condition  $c_1$  until message is not consumed, the sender component blocks all activities as it is in Figure 2;
- *post- blocking send*: it is similar pattern as before but sender component blocks all activities after satisfying post-condition  $c_2$  until  $c_3$ . This pattern is in Figure 3;
- *receive message*: in this pattern receiver blocks all activities until whole sent message is received;
- *lossy receive*: if the input port of receiver is not able to accept sent message, a message is lost.

The second kind of interactions concerns concurrency and cases of internal and external non determinisms. The difference between internal and external nondeterminism can be roughly specified as follows: internal non determinism does not depend on some decision or situation in environment, in external non determinism chosen activity depend on some external decision arising from environment. In this kind of interactions we distinguish the following patterns:

- *concurrent send (furcation)*: message is sent into two different receiver ports concurrently;
- *concurrent receive (rendezvous)*: two messages are receiving from two different ports concurrently,
- *sending choice (free choice)*: one message from two ready messages is sent according to choice of sender;
- *receiving choice (dependent choice)*: one of two sent messages is received depending on environment, it is external non determinism;
- *internal choice*: it is combination of the two possibilities above, the service does not make decision, neither the environment. This case is internal non determinism.

The third kind of interactions contains interactions based on what makes choice and what follows. In this kind we distinguish:

- *sending choice receiving follow-up*: the sender makes a choice and get information to the environment and then the environment send a message depending on the initial choice;
- *receiving choice sending follow-up*: the environment affects a choice of sender;
- *sending choice sending follow-up*;
- *receiving choice receiving follow-up*;
- *internal choice sending follow-up*.

For all these interactions above to be executed we have to use ports. But in real world there are situations when ports are faulty caused by containing some error, mismatching protocol or an incompatible data model. These faulty ports could cause shutting down interactions [17].

#### IV. TOWARDS FORMALIZATION OF COMPONENT COMPOSITION

In this section we present our preliminary ideas how formalize principles and properties of program systems constructed by composition of components.

After our good experience with several logical systems we assume that for description of basic principles can serve some logic with great expressive power. Using of logical system has another advantage, descriptions can be proved by corresponding deduction calculus. We choose linear logic [7] formulated by J.-Y. Girard. This logic can express dynamics and handling with such resources as time and space [22]. We discuss several properties of linear logic that are appropriate for describing basic principles of component based program systems.

Every formula in linear logic can be considered as action or a resource. Linear formulae are represented as types. That means, we can use formulae for denoting ports of components. Linear connectives have another useful properties:

- multiplicative conjunction  $A \otimes B$  expresses that actions  $A$  and  $B$  can be performed simultaneously;
- additive conjunction  $AB$  expresses dependent choice, external non determinism;
- multiplicative disjunction  $A \wp B$  expresses that if  $A$  is not perform then  $B$  is perform and vice versa;
- additive disjunction  $A \oplus B$  expresses free choice, internal non determinism;
- linear implication  $A \multimap B$  has special properties: a resource  $A$  is consumed after performing linear implication, and action  $B$  follows after  $A$ ;
- linear negation expresses consuming of a resource  $A$  or a reaction of action  $A$ .

Extending linear propositional logic with predicate symbols, linear term and quantifiers we get first order predicate linear logic with high expressive power. Contracts and dependencies can be express by predicates, i.e. relations. Similarly, all kinds of interactions are well formulable by linear formulae.

Sequent calculus of linear logic enables proving of constructed formulae and to provide a strong tool for verification purposes.

Formal description in linear logic has to be modeled in some appropriate structure. We prefer categories as very useful mathematical structures providing wide spectrum of interesting properties. In [10] is constructed categorical model for object oriented program system with some indication how to extend it for complex program systems.

Category theory provide two possibilities for modeling component base program systems [21][11]. If we are interesting in component composition then it is suitable to model component interfaces are category objects. In the case of modeling interactions as category morphisms we have two possibilities: either we can construct category morphisms as mappings expressing functionality of interactions or we can model them as relations which lead to relational categories. The latter approach is not obvious and it requires deeper analysis.

If we are interesting in modeling observable behavior of component program systems (for example [12]), it is suitable to use coalgebras [5], where polynomial endofunctor is constructed to model behavior of a system step by step. This endofunctor is constructed over category of states.

It is clear that the realization of our ideas requires much work in analysis, formal description together with verification and categorical modeling. Our aim is to prepare a formal system appropriately describing and interpreting basic principles of component based program systems.

#### V. CONCLUSION

Component based programming is an actual and rapidly evolving paradigm of programming. There are many useful research results in the frame of implementation, composition of components and eliminating arisen problems. We analyzed basic principles of this new method and formulated basic notions. Our analysis and discussion serves for our aim - to prepare a suitable formal system for verifiable formal description and modeling component based program systems.

#### ACKNOWLEDGMENT

This work has been supported by KEGA grant project No. 050TUKE-4/2012: "Application of Virtual Reality Technologies in Teaching Formal Methods", and by the Slovak Research and Development Agency under the contract No. APVV-0008-10: "Modelling, simulation and implementation of GPGPU-enabled architectures of high-throughput network security tools."

REFERENCES

- [1] AALST W.M.P van, MOOIJ A.J., STAHL C., WOLF K.: Service Interaction: Patterns, Formalization, and Analysis, In: *Proc. of 9th International School on Formal Methods for the Design of Computer, Communication, and Software Systems*, Bertinoro, Italy, 1-6 June 2009, LNCS Vol.5569, Springer, 2009, pp.42-88.
- [2] AALST W.M.P van der, HEE K.M. van, R.A. TOORN R.A van der: Component-Based Software Architectures: A Framework Based on Inheritance of Behavior, *Science of Computer Programming*, Vol. 42, No. 2-3, Elsevier, 2002, pp.129-171.
- [3] AOYAMA M.: New Age of Software Development: How Component- Based Software Engineering Changes the Way of Software Development, *International Workshop on CBSE, ICSE'98*, 1998.
- [4] BACHMANN F., BASS L., BUHMAN C., COMELLA-DORDA S, LONG F., ROBERT J., SEACORD R., WALL-NAU K.: *Technical Concepts of Component-Based Software Engineering*, Volume II, Carnegie Mellon, Techn.Report CMU/SEI-2000-TR-008, ESC-TR-2000-007, May 2000.
- [5] BARBOSA, L.: *Components as coalgebras*, PhD. Thesis, Uniuersity Minho, 2001.
- [6] COUNCILL B., HEINEMAN G.T.: Definition of a Software Component and Its Elements, Ch.1, *Component-based software engineering*, Addison-Wesley Longman, 2001, pp.5-19.
- [7] GIRARD, J.Y.: Linear logic, *Theoretical Computer Science*, Vol. 50, No.1, 1987, pp. 1-102.
- [8] HAN J.: *An Approach to Software Component Specification*, Melbourne, 2004.
- [9] HATCLIFF J., DENG W., DWYER M.B., JUNG G., RANGANATH V.: Cadena: An Integrated Development, Analysis, and Verification Environment for Componentbased Systems, *Software Engineering*, 2003, pp.160-172.
- [10] JENČÍK M., Mihályi D.: Category for component-based program system, *Electrical Engineering and Informatics*, Košice, 2012, pp. 575-579.
- [11] Knighten R.L.: *Notes on Category Theory*, MIT, 2011.
- [12] Harasthy, T., Turan, J., Ovsenik, L.: *Road line detection based on Optical Correlator*, In: *36th International Convention on Information & Communication Technology Electronics & Microelectronics (MIPRO)*, 2013, Opatija, 20-24 May 2013, pp. 298-300, IEEE 2013
- [13] Komponentenprogrammierung und Middleware- Component programming with C# and .NET, [http://www.dcl.hpi.uni-potsdam.de/LV/Components04/VL7/04\\_NET-components.pdf](http://www.dcl.hpi.uni-potsdam.de/LV/Components04/VL7/04_NET-components.pdf), April 2014.
- [14] KOZACZYNSKI W.: *Composite Nature of Component*, In: *International Workshop on Component- Based Software Engineering*, 1999 pp.73-77.
- [15] KWON O., YOON S., SHIN G.: Component-Based Development Environment: An Integrated Model of Object-Oriented Techniques and Other Technologies, In: *International Workshop on Component- Based Software Engineering*, 1999, pp. 2252-2256.
- [16] MAIN M., SAVITCH W.: *Data Structures and Other Objects Using C++*, Addison-Wesley Longman, 2010.
- [17] MARIANI L.: A Fault Taxonomy for Component-based Software, *Electronic Notes in Theoretical Computer Science*, Vol. 82, No. 6, 2003, pp. 55-65.
- [18] Meyer B.: Applying design by contract, *Computer*, Vol. 25, No.10, 2002, pp.40-51
- [19] Messabihi M., André P., Attiogbé C.: Multilevel Contracts for Trusted Components, In: *The 7th International Conference on Software Engineering Advances, ICSEA*, 2012, pp.71-85.
- [20] NING J.Q.: *A Component Model Proposal*, International Workshop on Component- Based Software Engineering, USA, 17-18 May 1999, pp.13-17.
- [21] NOVITZKÁ V., STEINGARTNER W.: *Categorical structures and their applications in informatics*, Equilibria, 2011 (in Slovak).
- [22] NOVITZKÁ V., MIHÁLYI D.: Resource-oriented programming based on linear logic, , Vol.4, No.2, 2007, pp.157-166.
- [23] POLBERGER D.: *Component technology in an embedded system*, Lund University, USA, 2009.
- [24] RAYMOND K.: *Reference Model of Open Distributed Processing (RM-ODP): Introduction*, University of Queensland, Brisbane, Australia, 1996
- [25] SZYPERSKI C., GRUNTZ D., MURER S.: *Component Software beyond Object- Oriented Programming*, ACM Press, New YORK, USA, 2002.
- [26] URTING D., BAELEN S. van, HOLVOET T.: Yolande BERBERS, *Embedded software development: components and contracts*, Belgium, 2001.
- [27] WANG A., Kai QIAN K.: *Component-oriented programming*, Wiley-Interscience, 2005.
- [28] WEGNER P.: *Concepts and Paradigms of object-oriented programming*, *ACM SIGPLAN OOPS*, Vol. 1 No. 1, 1990 pp. 7-87.
- [29] WILDE N., *Program Dependencies*, Carnegie Mellon, Tech.Report SEI-CM-26, 1990.
- [30] YACOUB S., AMMAR H., Ali MILI A.: *A Model for Classifying Component Interfaces*, ICSE'99 West Virginia, USA, May 1999.
- [31] ZOUHAR M: Quantification in natural language, *Philpapers*, 2008.

# Integration of new ICT solutions for Self-Government

*Anton TOMČAN, Miroslav MICHALKO*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

anton.tomcan@student.tuke.sk, miroslav.michalko@tuke.sk

**Abstract** — the aim of this work is to create an information system for higher territorial and local governments, specifically the municipality of Sabinov, but also for any organization. The most principal function of the system is the uniform creation, management and presentation of materials on activities of the municipality or any organization department to members or the public. The information system is represented by the web application called Veraz 1.0, which was created in PHP, HTML and JavaScript programming languages. The process of information saving during the processing of uniform materials by the user uses the MySQL database environment. The creation of the web application consisted of several parts from creation of database tables, through the forms for creation, editing and managing accompanying reports, the processing of source codes in a programming language to the final testing. The application offers comprehensive information on the created accompanying reports, minutes from the management, board and council meetings. The user interface is designed with the comprehensible graphical form using simple forms to create user-friendly web interface. The user has minimal text formatting options in order to maintain consistency of the output form of documents.

**Keywords** — web application, Form, PHP, MySQL, accompanying report

## I. INTRODUCTION

Public administration should focus its services to citizens and businesses. It should provide its services quicker and in a more efficient way. The vision for future is in cooperation between citizens and the state administration in order to make services more efficient, easier to understand and better accessible electronically.

The principal activities of the Municipal Office of the town Sabinov, as the executive body of the town council and the mayor, include the provision of a written agenda on town activities of self-government bodies and respective departments for members of the town council and the general public. The standard procedure is to create the accompanying report for the agenda item of the town council and town board meeting using relevant office packages. Since office packages offer a wide possibility to generate the output document even when using the existing template, the attainment of the uniform pattern of the presented document is unlikely. After internal printing, materials processed this way are transmitted to all Members of the Town Parliament for examination. The so-called corporate culture includes also the single written agenda, which is not always observed in the formation of these materials. The current requirements of Members of the municipal board include the electronic form of the accompanying reports with attachments.

The web application used for creating documents for local self-government with user and system manuals are also a part of the Work.

The main objective this Thesis is to create an online information system for unified document generation for local territorial government, management and editing of these documents, as well as enabling their online viewing by Members of the Town Parliament (MSZ) and general public, if necessary. The best user-friendly environment is the web interface using HTML forms. To create Veraz 1.0 web application it is necessary to perform several steps:

- create a design of the service and define its functionality;
- create the database of tables;
- create a user interface;
- create source codes of the web application.

The web application is intended to be used mainly by employees of municipal authorities whose role is to provide information on activities of respective departments. The system will also be used by members of the town board, who will have the opportunity to view the documentation before it is physically delivered. The expected benefit of the web application is to eliminate



printed copies of documents in gradual steps on condition such elimination is approved by the town management and MPZs.

## II. INFORMATIZATION OF PUBLIC ADMINISTRATION

In order to fulfil the strategic vision of public administration informatization, it is necessary to built eGovernment on a comprehensive approach that will coordinate, manage and control the process of eGovernment to a necessary extent. The conceptual approach will allow us to create preconditions to improve the provision of public services within the public administration to even up irregularities of the ICT implementation between the territorial self-government and state government. *"Computerisation/Informatization of public administration is a controlled process of creating the conditions for effective ICT application in public administration processes, as well as a controlled process of ICT application in public administration in the Slovak Republic."* [1].

The conditions necessary for correct ICT application in public administration include supporting new ICT applications in the public sector, efficient digitization of public administration sections, and last but not least satisfactory infrastructure. Informatization of public administration means the use of ICT capacities which assist in electronic performance of administration (e.g. Internet, mobile phone and other means of communication).

Positive reasons of computerization of public administration:

1. *The mission of public administration is to ensure legal requirements and rights of the public, and legal fulfilment of obligations of the public towards the state and the municipality;*
2. *Joining the EU, as well as increasing the competitiveness of Slovakia increase the number of interactions in the public sector, and thus demands on the performance of administration increase;*
3. *It is impossible to make difficult and qualified choices to promote the competitiveness of the country without up-to-date information processed in a timely manner* [1].

## III. ANALYSIS OF ICT CONDITION AT THE MUNICIPAL OFFICE OF SABINOV

Within the organizational structure of the Municipal Office, the Department of Informatics is included in the section of the Head of the Municipal Office. The role of the Department is to ensure main activities associated with the hardware functioning, its replacement, maintenance or potential repair. In the field of software, the Department manages information system of the office, APV Korwin, functioning of LAN and WIFI device, hot spot, e-mails, Internet access, installation and maintenance of purchased software products as well as the work related to the website of the Municipal Office [2].

Table 1  
Software tools of the Municipal Office (MO) of Sabinov - Operating Systems

Title	Number of licences
APV Korwin Information System	38 pcs
Microsoft Windows XP Professional CZ	43 pcs
Microsoft Windows Small Business Server 2003	1 pc
Linux – IBM System x3105	1 ks

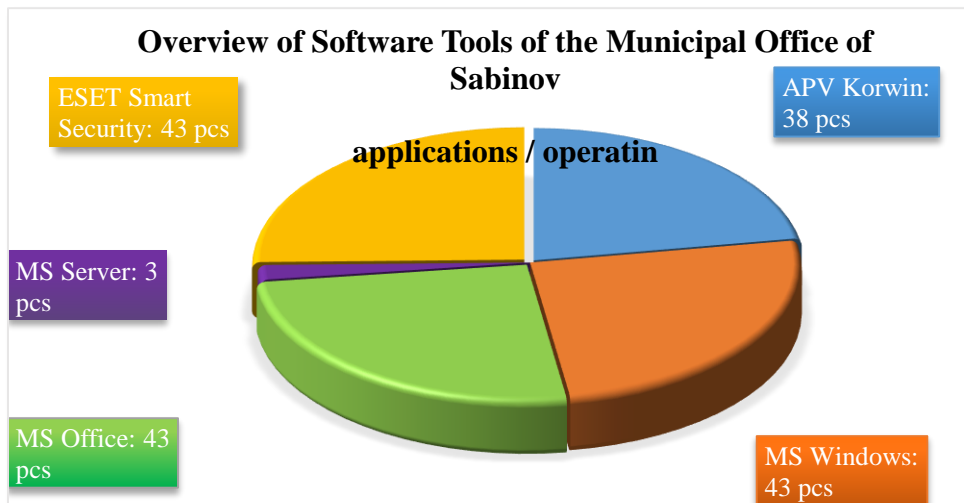


Fig. 1: Overview of software tools of MO Sabinov

#### IV. DESIGNING WEB APPLICATION THROUGH UML

In course of designing the web application, several types of charts were used to graphically present models of various aspects of the proposed application. From the behaviour diagrams, the activity diagram is used for application design (Fig. no. 2). The activity diagram shows respective transitions of the system for submission of the report accompanying the proposed agenda item of the Town Council / Town Board (TC/TB) from the proposal of the item by the supervisor, through the approval process, to the completion and delivery of materials to employees of the Office.

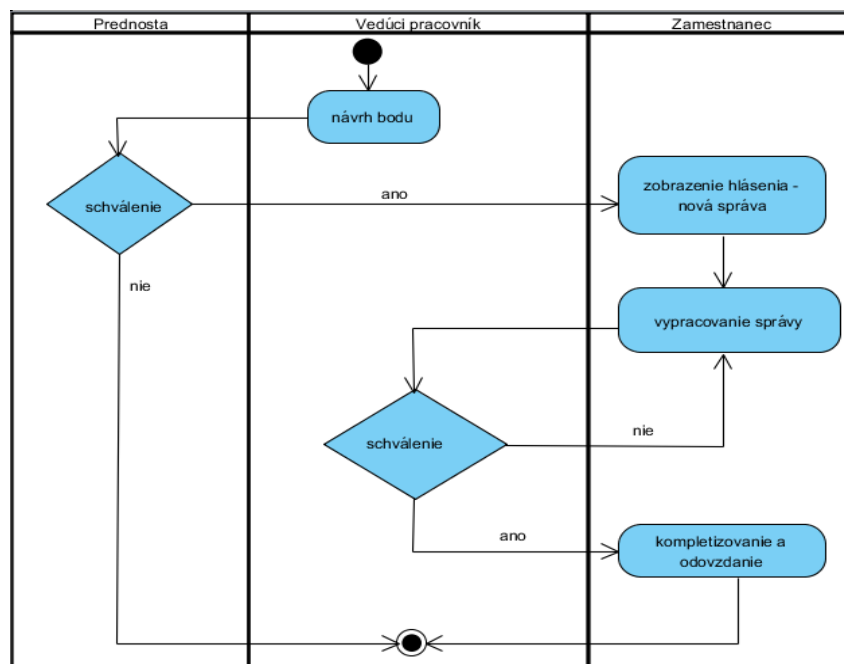


Fig. 2 - Activity diagram

The sequence diagram is used to view the individual steps in the process of chronological submission of materials for the TC/TB meeting. The component diagram shows the mutual interaction between the client that uses a web browser to obtain required information from the web application and to connect web and database servers which provide users with the required output information via PHP and HTML. Physical elements of the web application within the communication diagram are represented by respective user roles that share information to each other using the web environment on prepared accompanying reports, agenda items, and meeting minutes.

## V. ELECTRONISATION OF ACCOMPANYING REPORT

For members of the Town Board, the created web application represents the information system for easier access to presented materials in electronic versions. For Office staff, the web application is a tool to generate uniform accompanying reports in electronic form.

## VI. SERVICES AND FUNCTIONS OF VERAZ 1.0 WEB APPLICATION

The proposed web application defines multiple roles of user who are entitled to use the functions that the application provides. Administrator rights include especially standard handling of user profiles such as creation, editing and deleting.

Functional parts of the web application are very similar and reflect user roles according to their competence in the development of materials submitted.

Every full-time employee with the access to the corporate network may get familiar with minutes of management meetings and TC/TB meetings following the authorized login. The main activity of the employee is to create so called Matrix of the accompanying report and prepare materials for submission on the agenda item assigned to the employee. To draw up such accompanying report, the application offers an employee the form through which the employee may re-edit the material if necessary.

The main activity of the administrator is to manage user accounts within the web application which may be created, changed or the data of individual accounts may be modified or removed from the database, if necessary.

User interface is designed in the form of simple, forms easy to use even by less experienced users of web applications. The web application deployment diagram contains two nodes: the node for the server and the node for the client. The node on the part of the client is represented by the web browser that has to support JavaScript, CSS and HTML encryption in order to avoid the faulty display of a part of websites. To ensure client - server communication, server equipment with Windows Server 2012 R2 is used and it is located at the town hall.

## VII. DEPLOYMENT OF WEB APPLICATIONS AT THE MUNICIPAL OFFICE OF SABINOV

The structure of Veraz 1.0 database contains 9 tables. The first step of the implementation was to install database tables according to the scheme (Fig. no. 3) on the server of the municipal office.

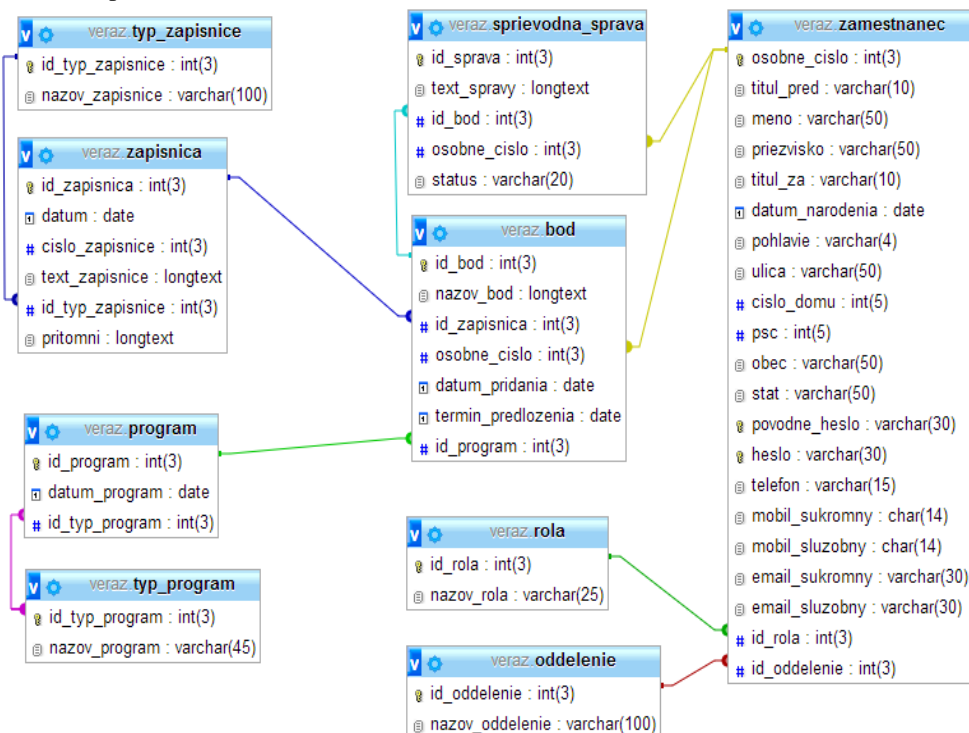


Fig. 3 Veraz Database Scheme

To log in the Veraz 1.0 web application login form is used which provides user authentication according to the role of the user at the Municipal Office.

The basic working environment consists of four parts - side information panel, user panel, info banner and the main information panel that provide users with all required information (Fig. 3). The proposed solution within Veraz 1.0 web application includes four user roles, which

are assigned an appropriate user panel by the application following the authentication. To create and edit accompanying reports by the user, the link "ACCOMPANYING REPORT" is created, which displays sent and approved / disapproved accompanying reports of the logged in user. To edit the selected accompanying report, the user may use a simple web form which consists of the text box and two buttons for save the report and finish the report handling. The Head of the Municipal Office updates the invitation to the meeting of the Town Council or the Town Board using the form to add the item in the agenda. After saving the item, new accompanying reports are displayed to individual employees in the "OUTBOX - NOT APPROVED". The administrator may use the form to modify user accounts and this form may be edited, if necessary.



Fig. 4 User interface of the web application

## VIII. WEB APPLICATION TESTING AT MO OF SABINOV

The heads of two departments, the head of the Municipality office and administrator, were asked to test the web application and they selected two employees from their department. For this purpose, the shared folder Veraz was established on the current server. The link to the index file of this folder was created at workstations of all concerned users. The personal number and password from the created list was assigned to every employee which can be used by the employee to log in the web application. Each employee is assigned their own personal number through which they can log in the system. All user accounts in the tested web application were created as imaginary in order to test the real deployment of the application. Individual users were continuously contacted in order to monitor the use of the web application and obtain the feedback.

## IX. CONCLUSION

In particular, the advantage of ICT use in public administration is to reduce costs of internal processes in creating information flows between departments at offices. The application of ICT in the public sector points out that ICT allows for more efficient work, and facilitates generation and control of documents for office staff.

The objective of the Thesis was to create the web application for material generation and editing for members of the Town Board to provide a short analysis of public administration, information and communication technologies and their interaction in Sabinov.

The created web application, Veraz 1.0, is the result of the Thesis. This application is primarily designed for employees of the Municipal Office. Members of the Town Board are the secondary users of this web application.

The web application is an online support for generation of uniform materials for Town Council or Town Board meetings. As a part of the user interface, the created application includes simple forms for material generation and editing. These forms were created also for inexperienced users. The application may also be modified for organizations which organize only management meetings. Electronisation of information provided to the Members of the Town Council and the Town Board, or even to ordinary citizens enables the self-government to ensure awareness and transparency of its activities towards natural persons and legal entities, or third sector.

## ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182,

supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is financed from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] Ministry of Finance of the Slovak Republic: Národná koncepcia informatizácie verejnej správy SR. [online, in SK language]. Bratislava: 2010. [cit. 2013-10-22]. Available online:<[http://www.informatizacia.sk/ext\\_dok-narodna-koncepcia-informatizacia-verejnej-spravy/6232c](http://www.informatizacia.sk/ext_dok-narodna-koncepcia-informatizacia-verejnej-spravy/6232c)> .
- [2] Bernát, L.: Analýza IT prostredia a návrh uplatňovania eGovernmentu v meste Sabinov [in SK language, Elektronická pošta]. Sabinov, Marec 2014.
- [3] Tomčan, A. Integration of new ICT solutions for self-government. Bachelor thesis. 2014. FEEL, Technical University of Kosice, Slovakia

# Ovládanie robotického ramena pomocou dátovej rukavice

<sup>1</sup>Peter GIROVSKÝ, <sup>2</sup>Radovan SIVÝ, <sup>3</sup>Matúš KUNDRÁT

<sup>1,2</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

<sup>1</sup>peter.girovsky@tuke.sk, <sup>2</sup>radovan.sivy@tuke.sk

**Abstrakt** — Tento príspevok sa zoberá návrhom ovládania robotického ramena pomocou dátovej rukavice. Použitá dátová rukavica bola zostrojená pomocou snímačov ohybu a snímača zrýchlenia. Použité robotické rameno bolo zostavené pomocou pohonov Dynamixel zo stavebnice Bioloid. Riadiaci program pre ovládanie robotického ramena pomocou dátovej rukavice bol zostavený v prostredí MATLAB.

**KLúčové slová** — robotické rameno, dátová rukavica, snímač ohybu, servopohon

## I. ÚVOD

Robotika má v dnešnom dobe svoje nenahraditeľné miesto najmä v oblasti priemyslu, kde si bez robotov už nevieme predstaviť žiadne automatizované pracovisko. Automatické roboty sú s výhodou používané kvôli vysokej rýchlosti a presnosti výroby. Robotika predstavuje komplexnú vednú oblasť, ktorá je úzko previazaná s elektronikou, elektrotechnikou, mechanikou, ako aj so softwarovým vybavením pre samotné riadenie robota. Prvý priemyselný robot bol nasadený do praxe už v roku 1961 firmou Unimation, kde jeho hlavnou úlohou bolo premiestňovať objekty. O tej doby robotika prešla obrovským vývojom a dnes sa s robotmi stretávame nielen na poli priemyselnej využiteľnosti, ale aj v iných oblastiach.

V poslednej dobe sa robotické systémy s ovládaním na diaľku stávajú jedným z najrozšírenejších druhov robotických systémov. Takéto robotické systémy nachádzajú svoje využitie v lekárstve, v záchranných akciách alebo v iných situáciách, kde sa namiesto ľudí posielajú do nebezpečného prostredia. Ovládanie takýchto robotických systémov na diaľku je možné pomocou rôznych spôsobov. Jeden z možných spôsobov je založený na snímaní pohybov človeka, ktorý takéhoto robota ovláda. Na snímanie pohybov existuje niekoľko systémov, pri ktorých sa využívajú napr.:

- infračervené kamery,
- magnetické systémy,
- mechanické systémy.

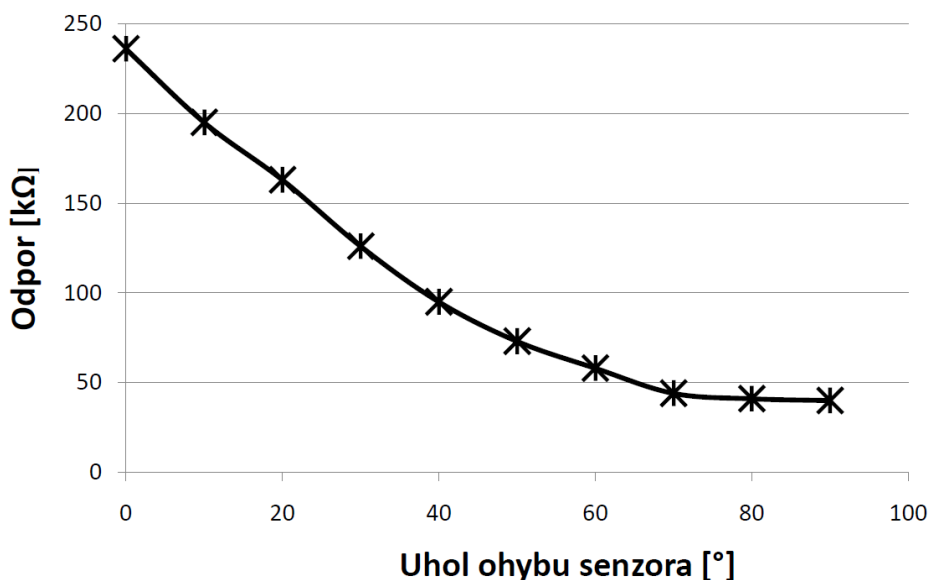
V tomto príspevok chceme predstaviť možnosť ovládania robotickej ruky pomocou snímania pohybu ľudskej ruky. Pre snímanie pohybov ruky bude použitá dátová rukavica. Na trhu existuje niekoľko typov dátových rukavíc, ktoré v podstate slúžia pre rovnaký účel. V našom prípade bola zostrojená dátová rukavica, v ktorej pre snímanie pohybu prstov slúžili snímače ohybu a pre snímanie pohybov celej ruky bol použitý akcelerometer.

## II. DÁTOVÁ RUKAVICA

Základom vyrobenej dátovej rukavice, zobrazenej na obr.2 sú snímače ohybu, uložené v prstoch rukavice, akcelerometer a riadiaca elektronika.

### A. Snímač ohybu

Pre našu dátovú rukavicu boli použité tri snímače ohybu, čo plne postačuje na ovládanie úchopovej časti robotickej ruky. Principiálne tento ohybový snímač predstavuje napäťový delič, kde odpor snímača, resp. výstupné napätie snímača (napäťového deliča), je úmerné uhlu ohybu senzora. Táto závislosť je zobrazená na obr. 1.



Obr. 1 Závislosť zmeny odporu od uhla ohybu snímača.

Výslednú hodnotu napätia  $U_s$  zo snímača, ktorá sa ďalej používa pre ovládanie polohy úchopovej časti robotickej ruky, môžeme určiť podľa nasledujúceho vzťahu (1), kde  $R_s$  predstavuje odpor snímača,  $R$  je prídavný odpor, a  $U$  je napájacie napätie 5V.

$$U_s = \frac{R_s}{R+R_s} \cdot U \quad (1)$$

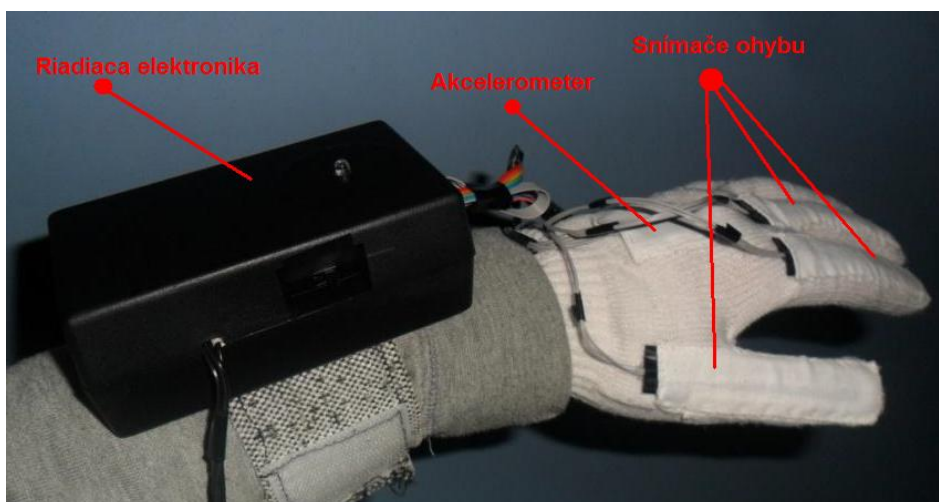
#### B. Akcelerometer

Pre snímanie náklonu ruky dopredu a dozadu, resp. doprava a doľava, sme zvolili akcelerometer MMA7361L.

Akcelerometer MMA7361L je trojosí akcelerometer od firmy Freescale s rozsahom zrýchlení 1,5g alebo 6g (v kladnom aj zápornom smere). Pri nepripojenom G-SEL pine je defaultne nastavené zrýchlenie 1,5g s citlivosťou 800 mV/g. Pri pripojenom pine je 6g s citlivosťou 206 mV/g. Pre snímanie pohybu ruky nám dostatočne postačuje zrýchlenie 1,5g. Pri tomto zrýchlení boli na akcelerometri zmerané hodnoty napätí vo všetkých troch osiach. Tieto hodnoty nám slúžili na riadenie pohybov celej robotickej ruky.

#### C. Riadiaci program dátovej rukavice

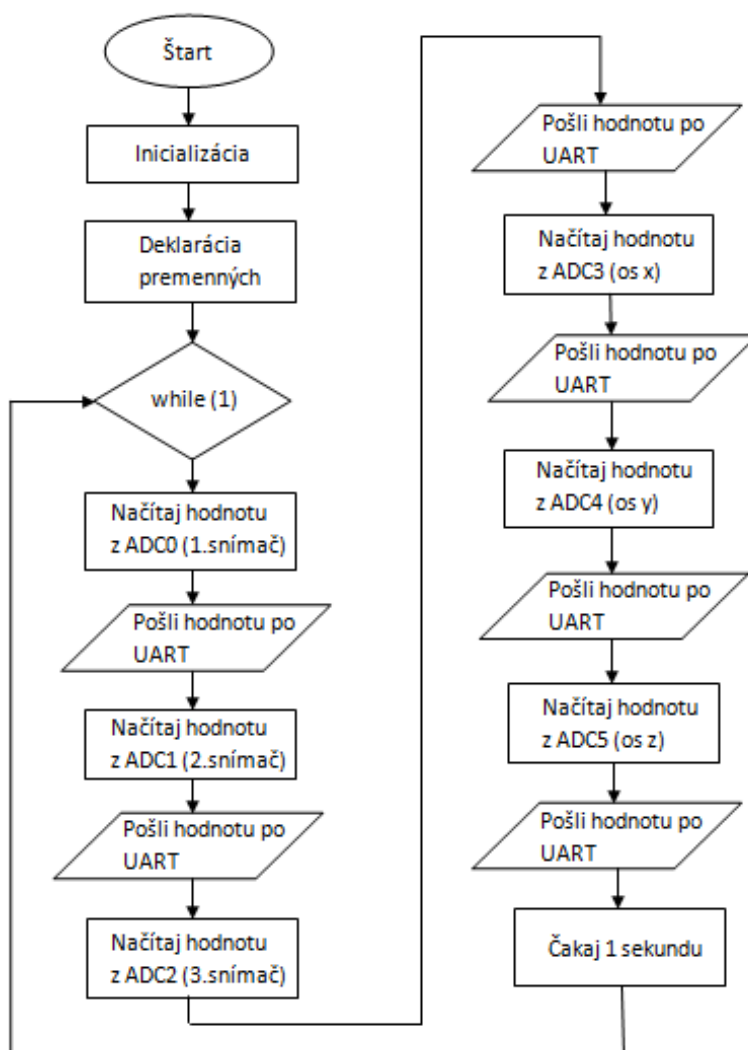
Pre riadenie celkového chodu dátovej rukavice, zobrazenej na obr.2 a komunikácie cez sériovú linku s programom Matlab, bol použitý mikrokontrolér ATmega8A.



Obr. 2 Dátová rukavica.

Principiálna schéma riadiaceho programu je zobrazená pomocou vývojového diagramu na obr.3.





Obr. 3 Vývojový diagram riadiaceho programu.

Pri inicializácii sa nastaví frekvencia mikrokontroléra a rýchlosť prenosu po sériovej linke. V programe *main* sa ďalej inicializuje ADC prevodník pre načítavanie hodnôt. Riadiaci program funguje tak, že v cykle *while* sa dookola načítavajú hodnoty z ADC portov, čiže hodnôt zo senzorov ohybu a z troch osí akcelerometra. Nakoniec sa každá hodnota pošle cez sériovú linku do počítača. Tento cyklus načítavania a posielania sa opakuje do nekonečna.

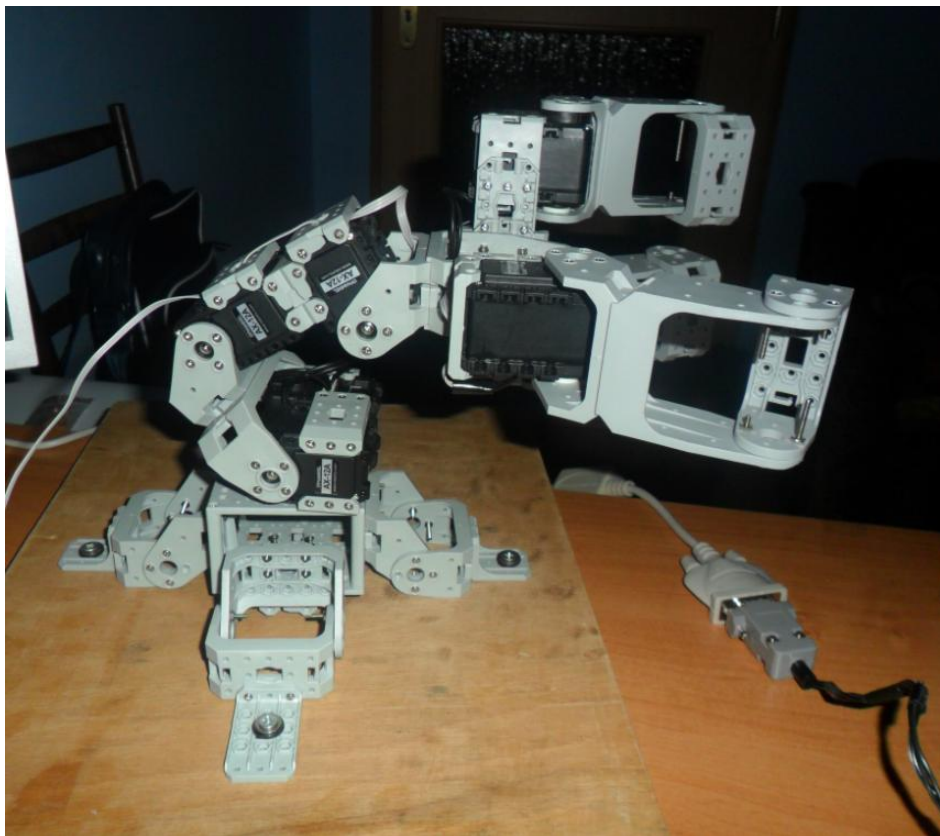
### III. OVLÁDANIE ROBOTICKEJ RUKY POMOCOU DÁTOVEJ RUKAVICE

Robotická ruka, zobrazená na obr. 4, bola zostrojená pomocou dielov stavebnice Bioloid Premium, kde pre pohyb boli použité servopohony Dynamixel AX-12A. Táto robotická ruka pozostávala zo siedmich servopohonov Dynamixel AX-12-A, z ktorých tri servopohony slúžili pre ovládanie úchopovej časti ruky, tri servopohony slúžili pre pohyb dopredu/dozadu a jeden servopohon pre otáčanie sa. Týmto spôsobom sem vedeli pohybovať s robotickou rukou v osiach *x*, *y*, *z* a ovládať uchopenie objektov.

Na riadenie a sledovanie pohonov Dynamixel AX12-A slúži tzv. riadiaca tabuľka. Používateľ môže riadiť Dynamixel tým, že zmení hodnotu inštrukcie v tabuľke. V našom prípade sme potrebovali ovládať parametre momentu, rýchlosti a aktuálnej polohy servopohonov.

Ovládať servopohony Dynamixel je možné rôznymi spôsobmi. V našom prípade bola robotická ruka a jej servopohony ovládané pomocou programu napísaného v jazyku C-mex v prostredí Matlab. Pre ovládanie robotickej ruky pomocou dátovej rukavice bolo potrebné napísať dva samostatné programy. Najskôr bol vytvorený program pre komunikáciu s dátovou rukavicou, kde prebiehalo načítavanie hodnôt z dátovej rukavice do Matlabu cez prevodník USB – RS232. V druhom kroku bolo potrebné vytvoriť program pre komunikáciu a ovládanie servomotorov dynamixel a pripojiť ho s programom pre získavanie informácií z dátovej rukavice.





Obr. 4 Robotická ruka.

#### A. Komunikácia s dátovou rukavicou

Komunikácia programu Matlab a dátovej rukavice prebiehala po sériovej linke. Z dátovej rukavice je posielaný spodný a horný byte pre každý snímač (3 snímače  $\times$  2=6) a pre každú os akcelerometra (3 osi  $\times$  2=6). Po načítaní hodnôt dátovej rukavice zo sériovej linky sa urobí prevod pre určenie polohy každého zo servopohonov. Keďže poloha každého zo servopohonov je určená v rozsahu 0-1023, prevodom dostávame hodnotu, ktorá je úmerná posunu servopohonu o jeden krok.

#### B. Komunikácia so servomohonmi Dynamixel

Pre komunikáciu so servomotorami Dynamixel je potrebné mať nainštalované knižnice Dynamixel SDK, ktoré obsahujú potrebné inštrukcie na ovládanie servomotorov Dynamixel. Samotná komunikácia medzi riadiaci programom v Matlabe a servopohonmi Dynamixel prebieha taktiež po sériovej linke.

Ak sa zmení hodnota napätia, prijatá z dátovej rukavice, zmení sa aj uhol natočenia servopohonu. Kvôli nepresnosti snímačov ohybu boli nastavené dve polohy uhla natočenia a to poloha, keď je ruka otvorená a poloha, keď je ruka zatvorená. Podprogram pre otáčanie bol riešený tak, že z prijímaných hodnôt je zadefinovaná hodnota, pri ktorej sa robotická ruka otáča doprava a hodnota, pri akom ide doľava. Taktiež bola definovaná stredová hodnota, čiže ak hodnoty pre os  $y$  sú v definovanom rozsahu tak sa robotická ruka nebude otáčať, ale ostal stáť. Pre pohyb dopredu a dozadu, boli zvolené koncové hodnoty, do ktorých sa môžu natáčať servomotory na základe hodnôt získaných z osi  $x$  akcelerometra.

## IV. ZÁVER

Hlavným cieľom tohto príspevku bolo predstaviť spôsob ovládania robotického ramena, zostrojeného z dielov stavebnice BIOLOID, pomocou dátovej rukavice. Samotné dátové rukavice sú pomerne drahou záležitosťou, preto pre naše riešenie bola zostrojená vlastná dátová rukavica. Pre riadenie dátovej rukavice bol použitý mikrokontrolér ATmega, ako snímače polohy prstov slúžili snímače ohybu a pre snímanie náklonu ruky v troch osiach bol použitý akcelrometer.

Ako už bolo spomenuté, samotná robotická ruka bola zostrojená z dielov stavebnice BIOLOID Premium. Hlavný riadiaci program, ktorý zabezpečoval komunikáciu a riadenie robotickéj ruky pomocou dátovej rukavice bol napísaný v programe Matlab. Pre ďalšie

vylepšenia tejto robotickej ruky by bolo vhodné využiť profesionálne snímače ohybu, ktoré sú presnejšie a taktiež využiť bezdrôtovú komunikáciu so zosťorenou robotickou rukou.

#### POĎAKOVANIE

Táto práca bola podporená projektom KEGA - 042TUKE-4/2012.

„Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: **Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií**, kód ITMS: **26220220182**, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.“



„Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ“

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.

#### LITERATÚRA

- [1] P. Karbanl.: Matlab a Simulink: Calculation and Simulation in the Programms. Brno: Computer Press vyd.2006. ISBN 80-251-1301-9 (in Czech).
- [2] Dynamixel AX-12 Manual. Available at: [http://www.robotis.com/xe/bioloid\\_en](http://www.robotis.com/xe/bioloid_en).
- [3] J. Bašista, Riadenie pohonov stavebnice BIOLOID v prostredí Matlab: Diplomová práca. Košice: TU FEI, 2013. 53 s.
- [4] M. Múdry, Dátová rukavica: Diplomová práca. Košice: TU FEI, 2012. 71 s.

# Odporúčací systém vytvorený v prostredí Apache Mahout

*Martin SARNOVSKÝ, Tomáš KURUC*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

`martin.sarnovsky@tuke.sk`, `tomas.kuruc@student.tuke.sk`

**Abstrakt** — Cieľom práce prezentovanej v tomto článku bolo navrhnuť a implementovať systém pre odporúčanie podobných položiek z databázy na základe podobnosti. Podobné prístupy sa využívajú napr. v internetových obchodoch, kde sa pri rastúcom objeme dát v databázach využívajú systémy pre spracovanie veľkých dát. Naším cieľom bolo navrhnuť a implementovať aplikáciu použitím takýchto systémov a technológií a otestovať ju na zvolenej dátovej kolekcii. Prezentovaný prístup je postavený na riešení využívajúcom rámec Apache Mahout pre metódy spracovania dát, NoSQL databáz a rámca Hadoop pre spracovanie veľkých objemov dát v distribuovanom prostredí.

**Kľúčové slová** — NoSQL, Hadoop, Mahout, Cloudové počítanie

## I. ANALÝZA VEĽKÝCH DÁT

Analýza veľkých dát, (Big data analysis) je analýza a spracovanie súborov dát, ktorých veľkosť je nad rámec schopnosti zachytávať, spravovať a spracovávať údaje bežne používanými softvérovými nástrojmi v rozumnom čase. Pojem "veľkosť dát" je chápaný nielen z hľadiska objemu dát, ale aj z hľadiska rýchlosti ich tvorby, prenosu a z hľadiska rôznorodosti ich typov. Ako príklad je často citovaných množstvo údajov o počasí, ktoré získava každý deň NOAA alebo NASA [1]. Veľké organizácie čelia stálej potrebe udržiavať rozsiahle súbory štruktúrovaných aj neštruktúrovaných dát. V súlade s vládnymi nariadeniami a s postupnou digitalizáciou narastá objem archivovaných elektronických dokumentov, e-mailových správ a ďalších rôznorodých záznamov o elektronickej komunikácii. Analýza veľkých dát je teda proces skúmania veľkých objemov dát za účelom odhaliť skryté vzťahy, neznáme korelácie a ďalšie užitočné informácie, ktoré môžu byť použité na efektívnejšie rozhodovanie. Vďaka technológiám umožňujúcim spracovanie veľkých dát je možné, že organizácie môžu zhromažďovať miliardy riadkov dát so stovkami miliónov kombinácií údajov vo viacerých dátových úložiskách bez ohľadu na formát [2]. Pre jednoduchšie a rýchlejšie spracovanie relevantných dát môžeme použiť vysoko výkonné analýzy. Použitie vysoko výkonného dolovania dát, prediktívnej analýzy umožňuje plynulo riadiť inovácie a vybrať čo najlepšie riešenie. Medzi analytické metódy patria prístupy strojového učenia, ktoré sú ideálne pre riešenie veľkých dátových súborov pomocou nových spôsobov.

Samotný pojem "Big data" môže byť dosť široký, podobne ako cloud, ktorý zahŕňa rôzne technológie [11]. Vstupné dáta do veľkých dátových systémov môžu byť konverzácie zo sociálnych sietí, logy webových serverov, satelitné mapy, vysielané zvukové streamy, bankové transakcie, MP3 – ky, obsahy webových stránok, skeny vládných dokumentov, GPS trás atď. Vo všeobecnosti je možné veľké dáta rozdeliť do troch rôznych skupín, podľa charakteristiky ich veľkosti [3].

### 1) Volume (Objem)

Mnoho spoločností vlastní veľké množstvo archivovaných dát, ale nemajú schopnosť ich spracovať. Za predpokladu, že objemy dát sú väčšie ako klasické relačné databázové infraštruktúry, možnosti spracovania sa rozdeľujú do voľby medzi masívne paralelné spracovanie architektúr dátových skladov alebo databáz.

### 2) Velocity (Rýchlosť)

Charakterizuje to zvýšenie rýchlosti, pri ktorej sa dátové toky dostávajú do organizácií, resp. ako dáta pribúdajú. Spracovanie rýchlo pribúdajúcich dát v reálnom čase je charakteristickou úlohou pri tomto stupni veľkosti dát. Medzi využívané prístupy patria streamovanie dát a

komplexné spracovanie udalostí. Kategórie produktov pre manipuláciu s dátovým tokom rozdeľujeme do zavedených patentovaných produktov, ako sú napríklad IBM InfoSphere Streams alebo stále novo vznikajúce open source nástrojov ako Twitter, Storm a Yahoo S4 [3].

### 3) *Variety (Rozmanitosť)*

Zriedka sú údaje prezentované v dokonalej podobe pripravené na spracovanie. Spoločným znakom vo veľkých dátových systémoch je, že zdroje dát sú rôznorodé a nepatria do úhladných relačných štruktúr. Mohli by to byť texty zo sociálnych sietí a obrázkové dáta. Žiaden z týchto príkladov nie je pripravený pre integráciu do aplikácií. Využívanie big data konceptov v tejto oblasti nám umožňuje použitie neštruktúrovaných dát a ich extrakciu do štruktúrovanej podoby ako vstupu pre aplikáciu.

## II. TECHNOLOGIE PRE SPRACOVANIE VEĽKÝCH DÁT

Hadoop je jeden z projektov Apache Software Foundation. Môžeme si predstaviť, že sa jedná o výpočtové prostredie postavené na distribuovanom súborovom systéme. Hadoop bol inšpirovaný prácou Googlu na vlastnom súborovom systéme (GFS, HDFS v Hadoope) a MapReduce. MapReduce nie je nová koncepcia, aj napriek tomu bol Hadoop prvý systém, ktorý ju urobil prakticky použiteľnou pre oveľa širšie zastúpenie úloh [13]. Hadoop je open-source softvér pre spoľahlivé, škálovateľné, distribuované výpočty. Umožňuje distribuované spracovanie veľkých dátových súborov cez klastre počítačov pomocou jednoduchých programovacích modelov. Je navrhnutý tak, aby škáloval jednotlivé servery na tisíce strojov, z ktorých každý ponúka lokálne výpočty a úložný priestor. Ako súborový systém Hadoop používa HDFS (Hadoop Distributed Filesystem) [12]. Má mnoho spoločného s existujúcimi distribuovanými systémami, ale rozdiely sú značné. HDFS je odolný voči chybám a je určený na zavedenie na nízkonákladovom hardvéri. Je vhodný pre aplikácie, ktoré obsahujú obrovské dátové súbory.

Ďalšou skupinou nástrojov používaných pre spracovanie veľkých dát sú NoSQL databázy. Je to názov, ktorý popisuje systém správy databáz, ktoré nepoužívajú klasický relačný model. Termín NoSQL nepopisuje presné technologické postupy a riešenia. Skôr zahŕňa všetky riešenia, ktoré nie sú založené na báze relačného modelu. NoSQL systémy pre správu databáz sú používané pri práci s veľkým množstvom dát, ktoré nemusia mať pevné usporiadanie. Dáta môžu byť štruktúrované, neštruktúrované alebo pološtruktúrované. V NoSQL sa využíva schopnosť uloženia a načítania veľkého množstva dát bez závislosti na vzťahoch medzi nimi [5]. Tento databázový systém sa vyvíjal počas vzniku obrovských internetových firiem, ako je Amazon, Facebook a Google, ktoré mali výrazne odlišné problémy zvládať enormné množstvo dát. Pre zabezpečenie plynulého a optimálneho výkonu sa dáta rozdeľujú. Obyčajne sú operácie obmedzené len na jednotlivé dátové položky. Systémy NoSQL pracujú s dátami, ktoré sú redundantne uložené na pár serveroch. Postupným pridávaním ďalších serverov je možné dosiahnuť jednoduchú škálovateľnosť systému, preto môžu byť prípadné zlyhania servera akceptované [4]. NoSQL databázové systémy sú často vysoko optimalizované pre pripojenie a načítanie operácií a obvykle ponúkajú iba veľmi málo funkcionality uloženého záznamu. Znížená časová flexibilita v porovnaní s SQL prístupom je kompenzovaná významným zvyšovaním škálovateľnosti a výkonu niektorých dátových modelov.

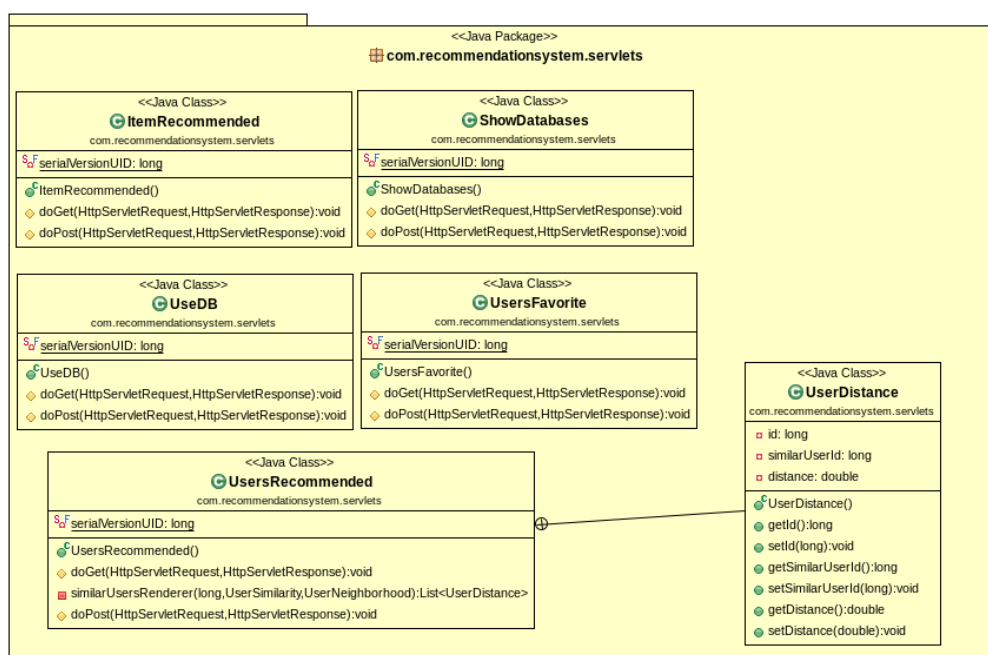
## III. NÁVRH A IMPLEMENTÁCIA APLIKÁCIE ODPORÚČACIEHO SYSTÉMU

Jednou z úloh, ktorá zápasí s problémom veľkých dát je odporúčanie položiek (napr. v internetových obchodoch) používateľom na základe podobnosti jednak položiek a aj samotných používateľov. V prípade rozsiahlych internetových obchodov disponujúcich s databázami ponúkaných produktov, ktoré z hľadiska veľkosti patria k veľkým dátam, predstavuje táto úloha typickú úlohu analýzy veľkých dát. Cieľom tejto práce bol návrh a vytvorenie aplikácie postavenej na technológiách pre spracovanie veľkých dát ako Hadoop, Mahout a NoSQL databázy pre riešenie takéhoto typu úlohy – odporúčacieho systému, ktorý generuje odporúčané položky v reálnom čase. Cieľom bolo implementovať systém bežiaci na lastri Hadoop, nad databázou NoSQL a využívajúci princípy paradigmy MapReduce za účelom zefektívnenia výpočtu. Front-end aplikácie predstavoval webovú aplikáciu s prezentačnou vrstvou.

Ako technológia NoSQL bola zvolená MongoDB. Medzi NoSQL databázami pravdepodobne najvyspelejšia dokumentovo orientovaná databáza a v niekoľkých aspektoch

je koncepcne blízka relačným databázam [6]. Každý záznam v databáze je chápaný ako dokument, ktorý má svoj vlastný identifikátor a operácie nad ním sú atomické. Dáta sú interne uložené a navonok prezentované vo formáte BSON, čo je špecifikácia knižnice pre prevod JSON do binárnej podoby. Databáza natívne podporuje veľké množstvo dátových typov, okrem tých bežných (celé aj neceločíselné číslo, reťazec neobmedzenej dĺžky, boolean, dátumočas), tiež odkazuje na iný dokument alebo binárne dáta. Dátový model je veľmi flexibilný a každé políčko môže obsahovať viac hodnôt a v políčku môže byť uložená ďalšia dátová štruktúra.

Pre analýzu dát bol zvolený rámec Mahout [7]. Je projektom spoločnosti Apache Software Foundation, ktorý je určený na produkciu voľných implementácií distribuovaných alebo inak škálovateľných algoritmov strojového učenia. Je zameraný predovšetkým na filtrovanie, zhľukovanie a klasifikáciu dátových súborov. Mahout tiež poskytuje Java knižnice pre bežné matematické operácie so zameraním na lineárnu algebru a štatistiku. Mahout je projekt, ktorý sa dá doplniť o nové funkcionality. Počet implementovaných algoritmov rýchlo narastá, ale množstvo algoritmov ešte stále nemá svoju implementáciu a čakajú na jej doplnenie.



Obr. 1 Diagram tried odporúčacieho systému.

Na získanie výpočtu odporúčení pre používateľa bolo použitých niekoľko základných metód, ktoré systém Mahout obsahuje ako Spearmanová, Euklidovská, Pearsonová a metóda City block.

### 1) Spearmanová metóda

Ide o neparametrickú metódu, ktorá sa využíva pri výpočte poradí hodnôt sledovaných veličín, nevyžaduje teda normalitu dát. Výhodou je, že sa dá táto metóda použiť pre popis akejkoľvek závislosti. Spearmanov korelačný koeficient používame najčastejšie pre meranie sily vzťahu u takých veličín, kde nemôžeme predpokladať linearitu očakávaného vzťahu alebo normálne rozdelenie sledovaných premenných. Závislosť premenných môže mať všeobecný vzostupný alebo zostupný ráz. Pre malé rozsahy je výpočet Spearmanovho korelačného koeficientu menej pracný než výpočet Pearsonovho korelačného koeficientu. Preto je možné ho použiť ako hodnotenie lineárnych závislostí. Jeho použitie je tu však skôr orientačné na rozdiel od parametrického koeficientu je menej účinné [8].

### 2) Euklidovská metóda

Podobnosť medzi dátovými objektmi môže byť zastúpená rôznymi spôsobmi. Prvý spôsob, ktorý možno použiť je vzdialenosť. Vzhľadom k tomu, že dátové objekty sú abstraktné pojmy vzdialenosť medzi nimi je súčtom vzdialeností jednotlivých atribútov dátových objektov.

Najjednoduchší spôsob výpočtu podobnosti dosiahneme použitím Euklidovskej metódy, ktorá predstavuje dĺžku prepony pravouhlého trojuholníka a jej výpočet je založený na Pytagorovej vete.

### 3) *Pearsonová metóda*

Na rozdiel od Euklidovskej metódy na meranie podobnosti, ktorá meria hodnoty od 0 do 1, je táto metóda vysoko korelovaná a premenné sú merané v rozmedzí -1 po +1. Podobne ako u modifikovaných Euklidovských vzdialeností, keď je Pearsonov korelačný koeficient je rovný +1, znamená to, že dátové objekty sú perfektne korelované. V prípade, že hodnota korelačného koeficienta je -1 znamená to, že dátové objekty majú najvyššiu negatívnu koreláciu. Hodnota 0 vypovedá žiadnu koreláciu. Najviac používaným typom korelačného koeficienta je Pearsonov, ktorý sa používa, ak sú premenné merané prinajmenšom na intervalovej škále. Korelačný koeficient nezávisí od mierky, v ktorej boli premenné merané. Môžeme korelovať výšku s hmotnosťou, ktoré sú vyjadrené v centimetroch a kilogramoch. Korelačný koeficient vyjde rovnaký. Korelácia je vysoká, ak sa dá meranými bodmi v rovine metódou najmenších štvorcov "dobře preložit" priamka. Metóda najmenších štvorcov znamená, že súčet umocnených vzdialeností meraných bodov od regresnej priamky je najmenší [9].

### 4) *Metóda city block*

Tento typ metriky sa nazýva aj manhattanská. Ak pracujeme s dvojrozmernými pozorovaniami ide o vzdialenosť dvoch bodov v rovine, ktorá je meraná na odvesnách pravouhlého trojuholníka, zatiaľ čo Euklidovská vzdialenosť je meraná na prepone. Určuje sa ako suma vzdialeností všetkých párov prvkov oboch pozorovaných vektorov. Manhattanská metóda vzdialenosti sa používa pre ordinálne premenné, to znamená pre znaky, ktoré možno zoradiť podľa veľkosti. Odporúča sa aj v prípadoch znakov, ktoré nespĺňajú predpoklad normality [10].

Diagram tried (Obr. 1) popisuje štruktúru systému zobrazením tried a vzťahov medzi nimi. V tejto kapitole si v krátkosti opíšeme funkcionality použitých tried. V triede *ItemRecommended* prebieha výpočet odporúčaní pre zvolenú položku na základe jednotlivých metód z knižnice Mahout, ktoré sme použili pre výpočet podobnosti. Trieda si vytvára objekt na výpočet odporúčaní z generickej triedy a vyberá 10 odporúčaných položiek a následne odošle tento zoznam položiek na stránku na zobrazenie. Servlet slúži na vytvorenie zoznamu databáz obsahujúcich údaje na výpočet odporúčaní. *ShowDatabases*, načíta zoznam databáz a následne ho zobrazí vo výberovom formulári pre používateľa. Trieda obsahuje výber databázového pripojenia zo session. Následné vrátenie zoznamu databáz zmaže trieda predvolené databázy zo zoznamu. Nakoniec samotný zoznam nastaví ako atribút servletu pre zobrazenie na stránke. V prípade akejkoľvek chyby, najčastejšie vypršanej session, nás vráti na úvodnú obrazovku. *UseDB* na základe zvolenej databázy vytvorí dátový model pre výpočet podobností a odporúčaní, a to nasledujúcim spôsobom. V prvom kroku si z URL parametra vyberieme názov zvolenej databázy a zo session vyberáme pripojenie na databázu, ktorej je potrebné nastaviť pripojenie. Samotný výpočet dátového modelu slúži na ďalšie výpočty odporúčaní. Zvolenú databázu použijeme na výpis používateľov, kde je v prvom rade potrebné zvoliť zoznam používateľov a nastaviť podmienku pre výber používateľov. Následne je tento zoznam odoslaný na zobrazenie na stránke. V triede *UsersFavourite* si z URL vyberáme ID zvoleného používateľa a zo session vyberieme pripojenie na databázu. Následne vyberieme zvolenú databázu a zvolíme zoznam hodnotení používateľov. Nastavíme podmienky pre zobrazenie a zoradenie výsledkov. Vyberieme údaje na základe podmienok a zoznam hodnotení odošleme na stránku na zobrazenie. Z URL vyberáme hodnotu parametra zo zvoleným používateľom a nastavujeme konfiguráciu logovania pre Mahout, ktorú vypíšeme na konzolu. V tejto triede je obsiahnutý výpočet odporúčaní pre používateľa na základe jednotlivých metód, ktorý je zložený z viacerých bodov. Prvým bodom výpočtu je výpočet na základe zvolenej metódy. Ďalej je tu možnosť výberu 20 najbližších používateľov na základe vypočítanej podobnosti. Bol vytvorený cache objekt z predošlého výpočtu kvôli zrýchleniu nasledujúceho. Trieda ponúka možnosť výberu 10 odporúčaných položiek a odoslanie zoznamov odporúčaných položiek a podobných používateľov na zobrazovaciu stránku. Nasleduje samotné volanie stránky so zobrazením odporúčaných položiek a podobných používateľov. Zoznam miery podobnosti a používateľov je vytvorený funkciou. Táto trieda obsahuje pomocnú triedu na zobrazenie

podobnosti medzi dvoma používateľmi.

#### IV. VYHODNOTENIE

Ako dáta na testovanie aplikácie sme použili dáta movielens<sup>1</sup>, ktoré sú štandardnou databázou pre tento typ úloh. Obsahuje databázu filmov, ich hodnotení používateľmi a samotných používateľov. Pozostáva z atribútov ako ID používateľa, ID filmu, žáner filmu, jednotlivé hodnotenia používateľov, vek, pohlavie a mnoho iných. Dátové súbory boli zhromaždené v priebehu rôznych časových úsekov a v závislosti na veľkosti súboru.

Na testovanie aplikácie boli použité dáta s názvom MovieLens 1M, ktoré pozostávajú z 1 000 209 anonymných hodnotení, 3900 filmov a 6040 používateľov, ktorí hodnotili dané položky. Cieľom bolo prostredníctvom webovej aplikácie otestovať funkčnosť odporúčacieho systému. Testovacia webová aplikácia pozostáva z niekoľkých rozhraní pre nastavenie parametrov riešenej úlohy a z rozhrania pre import testovacej databázy. Výstupom aplikácie bolo rozhranie, ktoré vizualizuje vypočítané položky na zadaný dotaz (napr. odporúčané filmy podľa podobnosti ku konkrétnemu filmu, alebo odporúčané filmy na základe podobnosti používateľov) podľa jednotlivých metód. Cieľom nebolo skúmať samotné metódy, ale funkčnosť a integráciu používaných technológií pre takýto typ úlohy. Na obr. 2 je zobrazený výstup aplikácie – zoznam odporúčaných produktov podľa troch testovaných metrick podobnosti.

Recommended items for Broken Arrow

Pearson Method	
Movie Title	Rating
<a href="#">Dracula: Dead and Loving It</a>	0.56067216
<a href="#">Wings of Courage</a>	0.52304405
<a href="#">Bed of Roses</a>	0.49854887
<a href="#">Nico Icon</a>	0.4695626
<a href="#">Friday</a>	0.4444988
<a href="#">Restoration</a>	0.42571613
<a href="#">Screamers</a>	0.39062673
<a href="#">How to Make an American Outhouse</a>	0.37628454
<a href="#">Othello</a>	0.37550813
<a href="#">Mr. Holland's Opus</a>	0.37475964

Euclidean Method	
Movie Title	Rating
<a href="#">Wings of Courage</a>	0.42823917
<a href="#">Othello</a>	0.4258219
<a href="#">Lamerica</a>	0.41190454
<a href="#">Restoration</a>	0.40557155
<a href="#">Dracula: Dead and Loving It</a>	0.39911473
<a href="#">Bed of Roses</a>	0.3981117
<a href="#">Mr. Holland's Opus</a>	0.39696574
<a href="#">Nico Icon</a>	0.394989
<a href="#">French Twist (Gazon maudit)</a>	0.3908712
<a href="#">Georgia</a>	0.3874259

Uncentered Cosine Method	
Movie Title	Rating
<a href="#">Wings of Courage</a>	0.9300397
<a href="#">Othello</a>	0.9216688
<a href="#">Dracula: Dead and Loving It</a>	0.9195015
<a href="#">French Twist (Gazon maudit)</a>	0.9131555
<a href="#">Lamerica</a>	0.9112964
<a href="#">Restoration</a>	0.91092783
<a href="#">Nico Icon</a>	0.90092224
<a href="#">How to Make an American Outhouse</a>	0.90068746
<a href="#">Bed of Roses</a>	0.89829475

Obr. 2 Zoznam odporúčaných produktov vypočítaných na základe podobnosti používateľov.

#### V. ZÁVER

Cieľom tejto práce bolo vytvorenie aplikácie pre odporúčací systém používajúci technológie veľkých dát a k nej rozhranie vo forme webovej aplikácie. Pri vytváraní samotnej aplikácie, boli použité rôzne technológie, ktoré slúžia na prácu s veľkými súbormi dát, ako Hadoop, Mahout, MapReduce a MongoDB. Aplikácia bola otestovaná v testovacom pseudo-distribúovanom prostredí, ktoré simulovalo jej nasadenie na reálnom klastri, ktorý sme v šase testovania nemali k dispozícii. Základná štruktúra back-endu aplikácie môže byť použitá pre riešenie iných odporúčacích úloh, alebo pre integráciu do reálneho internetového portálu v ktorom je možné využiť tento typ riešenej úlohy.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na

<sup>1</sup> <http://grouplens.org/datasets/movielens/>



základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### REFERENCIE

- [1] O. Dolák, Big data Nové způsoby zpracování a analýzy. [online] Publikované 18.01.2011. [citované 16.03.2014] dostupné z < <http://www.systemonline.cz/clanky/big-data.htm> >.
- [2] K. Collins, Big Data Analytics. [online] Publikované 8.01.2013. [citované 16.02.2014] Dostupné z < [http://www.sas.com/en\\_us/insights/analytics/big-data-analytics.html](http://www.sas.com/en_us/insights/analytics/big-data-analytics.html) >.
- [3] E. Dumbill, What is big data. [online] Publikované 11.01.2012. [citované 13.02.2014] Dostupné z < <http://strata.oreilly.com/2012/01/what-is-big-data.html> >.
- [4] S. Tiwari, Professional NoSQL 1. vyd. Indianapolis: John Wiley & Sons, 2011. 384 s. ISBN: 978-0-470-94224-6
- [5] C. Strauch, U. Sites, W. Kriha: "Nosql databases," Lecture Notes, Stuttgart Media University, 2011
- [6] T. Hawkins, E. Plugge, P. Membrey, Peter: The Definitive Guide to MongoDB: The NoSQL Database for Cloud and Desktop Computing (1st ed.), Apress, p. 350, ISBN 978-1-4302-3051-9
- [7] S. Owen, R. Anil, T. Dunning, E. Friedman, Mahout In Action, Manning Publications Co., Manning Publications Co. 20 Baldwin Road PO Box 261 Shelter Island, NY 11964, First edition, 2011
- [8] J.L. Myers, A. D. Well, Research Design and Statistical Analysis (2nd ed.). Lawrence Erlbaum. p. 508. ISBN 0-8058-4037-0, 2003
- [9] R. R. Wilcox, Introduction to robust estimation and hypothesis testing. Academic Press, 2005.
- [10] E. F. Krause, Taxicab Geometry. Dover. ISBN 0-486-25202-7, 1987
- [11] F. van der Molen, Get ready for Cloud Computing. 1. vyd. Leuven: Van Haren Publishing, ISBN 978 90 8753 640 4, 2010
- [12] K. Shvachko, H. Kuang, S. Radia, R. Chansler, The Hadoop Distributed File System, In Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies, IEEE Computer Society, Washington, DC, USA, 2010
- [13] T. White, Hadoop: The Definitive Guide (1st ed.). O'Reilly Media, Inc., 2009.



# Využitie DCT a SVD pri návrhoch algoritmov digitálnej vodotlače

Gabriel BUGÁR, Vladimír BÁNOCI, Dušan LEVICKÝ, Michal DUDÁŠ

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská Republika

{gabriel.bugar, vladimir.banoci, dusan.levicky}@tuke.sk,  
michal.dudas.3@student.tuke.sk

**Abstrakt** — Tento článok je zameraný na porovnanie dvoch často používaných transformácií pri návrhoch algoritmov digitálnej vodotlače. Konkrétne budú rozobrané transformácie DCT a SVD, z hľadiska ich využitia pri tvorbe algoritmov steganografie resp. digitálnej vodotlače. V závere sú zhrnuté výsledky dvoch základných algoritmov pre obe transformácie.

**KLúčové slová** — digitálna vodotlač, diskrétna kosínusová transformácia, singular value decomposition,

## I. ÚVOD

V dnešnej modernej dobe, v ktorej sme obkolesení Internetom a tým aj útočníkom či iným skupinám, ktorým cieľom je obohatiť sa na majetku iných je dôležité zlepšovať metódy ich chránenia si. Jednou z týchto oblastí je aj digitálna vodotlač. Aj keď v táto oblasť je dlhé roky skúmaná a vylepšovaná stále je priestor na jej zdokonaľovanie a hľadanie nových, lepších metód utajovania obsahu resp. ochrany dát. Táto práca je zameraná práve na jednu z metód digitálnej vodotlače v transformovanej oblasti v ktorej má hlavné zastúpenie transformácie DCT a DWT. Preto sa hľadali nové transformácie, ktoré by vylepšili ich nedostatky alebo poprípade v spojení s inými dosiahli lepšie výsledky. V ďalších kapitolách bude rozobraná transformácia SVD (singular value decomposition) ako ďalší z možných nástrojov na vylepšovanie metód digitálnej vodotlače a zároveň porovnaná s asi najčastejšie doposiaľ využívanou transformáciou DCT.

## II. ZÁKLADY DIGITÁLNEJ VODOTLAČE

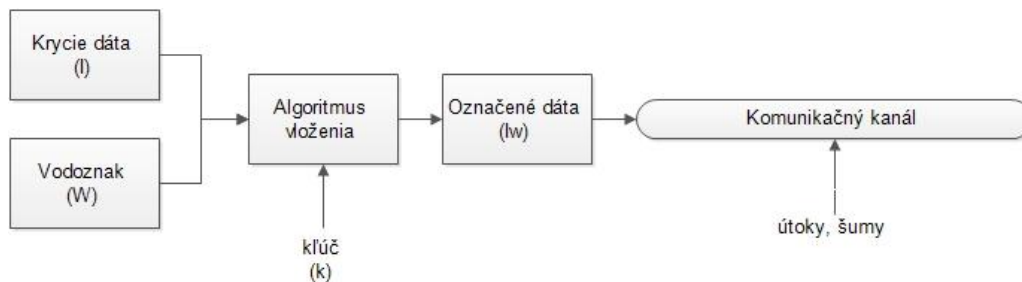
Digitálna vodotlač je využívaná na vloženie prídavnej informácie (vodoznaku) do multimediálnych dát tak, aby modifikácia týchto dát bola zmyslovo nepostrehnuteľná a aby bolo možné túto vloženú informáciu extrahovať len na základe znalosti tajného, resp. verejného kľúča. Prídavná informácia, ktorá je vkladaná do multimediálnych dát, zvyčajne nesie informácie o autorovi, o autorizovanom príjemcovi týchto dát alebo informácie o kópii multimediálnych dát. Digitálna vodotlač nachádza uplatnenie hlavne v aplikáciách autentizácie a ochrany autorských práv multimediálnych dát [1][2].

DRM (digital right management) je práve jedna z oblastí v ktorej sa uplatňujú metódy digitálnej vodotlače. DRM môžeme považovať za skupinu technických metód a nástrojov na zabezpečenie ochrany autorských práv. Práve kvôli tomuto majú aj algoritmy digitálnej vodotlače na rozdiel od klasických steganografických algoritmov zvýšený nárok na robustnosť okrem všetkých ostatných základných požiadaviek a síce: nedetekovateľnosť (*invisibility*), nevnímateľnosť (*undetectability*), bezpečnosť (*security*), kapacita (*capacity*), robustnosť (*robustness*), bezpečnosť (*security*).

Každá z týchto požiadaviek má svoje vlastné opodstatnenie a dôležitosť v algoritme a práve na robustnosť je v digitálnej vodotlači dávany najväčší dôraz.

Vo všeobecnosti vieme každý algoritmus opísať pomocou Obr. 1, kde vstupom sú krycie dáta „I“ v našom prípade krycí obraz a vodoznak „W“ čiže tajná správa, ktorá bude do krycích dát vložená. Nasleduje algoritmus vloženia, ktorý poznajú obe strany poprípade aj tajný kľúč „k“, ktorý slúži na správne vyextrahovanie vodoznaku. Takto označené krycie dáta putujú komunikačným kanálom, na ktoré môžu a nemusia byť v procese prenosu poškodené. Práve preto má digitálna vodotlač zvýšené nároky na robustnosť aby v prípade cieľového útoku alebo aj pôsobením šumov a chýb v prenosovom kanále bol možný adresát resp. vlastník dát vyextrahovať daný vodoznak. V súčasnosti existuje široká škála možných útokov na

vodoznaky pričom cieľom tzv. robustného vodoznaku je zabezpečiť čo možno najväčšiu odolnosť voči týmto rôznym útokom. Na strane adresáta nasleduje inverzný proces (extrakcia).



Obr. 1 Schéma vloženie

### III. DEKOMPOZÍCIA NA BÁZE SVD

Princípom dekompozície obrazu resp. akejkoľvek matice  $A$  o rozmeru  $M \times N$  a hodnoty  $r$  pomocou SVD je, že ju dokážeme vyjadriť ako: **Error! Reference source not found.**[4]

$$A = UxSxV^T \quad (1)$$



Obr. 2 Rozklad matice A pomocou SVD

pričom platí, že  $r \leq m \leq n$ . Matica  $U$  je ortogonálna unitárna matica rozmeru  $M \times M$ :

$$U = [u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_m] \quad (2)$$

ktorú tvorí set  $i=1$  až  $m$  ortonormálnych stĺpcov  $u_i$ :

$$u_i^T \cdot u_j = \delta_{ij} = f(x) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (3)$$

Podobne platí pre maticu  $V$  rozmeru  $N \times N$ , že je tvorená  $i=1$  až  $n$  ortonormálnymi stĺpcami  $v_i$ . Stĺpcové vektory  $u_i$  a  $v_i$  sa tiež zvyknú nazývať ľavé a pravé singulárne vektory matice  $A$ . Matica  $S$  je diagonálna matica rozmeru  $N \times N$ , ktorá je tvorená singulárnymi hodnotami  $\sigma_i$  na diagonále a zvyšné hodnoty sú rovné nule. Pre singulárne hodnoty  $\sigma_i$  platí, že:

$$\begin{cases} \sigma_1 \geq \sigma_2 \geq \sigma_3 \dots \geq \sigma_r > 0 \\ \sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0 \end{cases} \quad (4)$$

$$S = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \sigma_{r+1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}$$

Obr. 3 Singulárne hodnoty SVD

Vidíme, že SVD nie je klasickou transformáciou ako napr. DCT alebo DFT, ktoré transformujú obraz z časovej do frekvenčnej oblasti. SVD je matematickým nástrojom, ktorý pomocou určených pravidiel za pomoci troch matíc opíše geometrické vlastnosti obrazu podobne ako spomínané iné transformácie.

*A. Teorémy o SVD*

*1) Existencia SVD*

Ak  $A \in \mathbb{R}^{m \times n}$ , tak existujú aj matice U a V a platí:

$$U^T \cdot A \cdot V = S \tag{5}$$

pričom S pozostáva zo singulárnych hodnôt  $\sigma_i$ , ktoré reprezentujú „eigenvalues“  $\lambda_i$   $A \cdot A^T$  resp.  $A^T \cdot A$  a síce  $\sigma_i = \sqrt{\lambda_i}$ .

*2) Stabilita SVD (The stability of SVD)*

Stabilita SVD hovorí o tom, že ak nastala malé porušenie A, zmena jej singulárnych hodnôt nebude väčšia než 2-norm matice ktorá ju narušila. 2-norm predstavuje najväčšiu singulárnu hodnotu. Vo všeobecnosti môžeme povedať, že zmena matice A len málo narúša singulárne hodnoty matice S.

*3) Škálovanie SVD (The scaling property)*

Ak singulárne hodnoty matice  $A^{m \times n}$  sú rovné  $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_n$  tak pre singulárne hodnoty matice  $\beta \cdot A^{m \times n}$   $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_n$  platí :

$$|\beta|(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_n) = (\sigma_1', \sigma_2', \sigma_3', \dots, \sigma_n') \tag{6}$$

*4) Nezávislosť na otočení (The rotation invariant property)*

Ak P je unitárna a rotačná matica (rotating matrix), singulárne hodnoty matice PA (otočená) budú rovnaké ako singulárne hodnoty matice A.

*5) Nezávislosť na preusporiadaní (The translation invariance property)*

Originálny obraz A a obraz A' ktorý je tvorený tými istými preusporiadanými riadkami a stĺpcami má rovnaké singulárne hodnoty ako matica A.

*6) Nezávislosť na transpozícii*

Ak

$$A \cdot A^T \cdot u = \lambda^2 \cdot u \tag{7}$$

potom

$$A^T \cdot A \cdot v = \lambda^2 \cdot v \tag{8}$$

a síce, A a  $A^T$  majú rovnaké singulárne hodnoty.

#### IV. DCT

Pri technikách využívajúcich DCT sa zvyčajne originálny obraz rozdelí na bloky, ktoré sú pomocou DCT transformované postupne. To znamená, že DCT transformuje každý obrazový blok do bloku transformačných koeficientov, pozostávajúceho z jednosmerného koeficientu DC a striedavých koeficientov AC1 až AC63. Na Obr. 4 je znázornený blok DCT koeficientov. Koeficient v ľavom hornom rohu je jednosmerný koeficient (DC) predstavujúci priemernú hodnotu šedej úrovne celého obrazového bloku. Ďalšie prvky v bloku sú AC koeficienty, ktoré zodpovedajú nenulovým priestorovým frekvenciám a stúpajú v zmysle sekvencie, ktorá sa označuje "cik-cak"[6].

DC	1	5	6	13	14	26	27
2	4	7	12	15	25	28	42
3	8	11	16	24	29	41	43
9	10	17	23	30	40	44	53
19	18	22	31	39	45	52	54
20	21	32	38	46	51	55	60
34	33	37	47	50	56	59	61
35	36	48	49	57	58	62	63

**Obr. 4 Rozloženie koeficientov v bloku 8x8**

DCT transformácia je jednou z najčastejšie používanou transformáciou kvôli jej ľahkému výpočtu (rovnice 9 a 10) taktiež fakt, že DCT tvorí základ štandardu JPEG a preto dosahuje aj dobré výsledky voči kompresným útokom.

$$F(u, v) = \frac{C(u)C(v)}{4} \cdot \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cdot \cos\left[\frac{(2i+1)u\pi}{16}\right] \cdot \cos\left[\frac{(2j+1)v\pi}{16}\right] \quad (9)$$

$$f(i, j) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(u, v) \cos\left[\frac{(2i+1)u\pi}{16}\right] \cdot \cos\left[\frac{(2j+1)v\pi}{16}\right] \quad (10)$$

$$C(u) = 1/\sqrt{2} \text{ pre } u=0 \text{ a } 1 \text{ pre } u=1,2,\dots,7$$

$$C(v) = 1/\sqrt{2} \text{ pre } v=0 \text{ a } 1 \text{ pre } v=1,2,\dots,7$$

#### V. POROVNANIE METÓD DCT A SVD

V tejto práci boli využité dve metódy každá pre jeden typ transformácie. V prípade metódy s využitím SVD to bola metóda priamej modifikácie singulárnych koeficientov a v prípade DCT metóda vloženie na základe zámene koeficientov na základe určitej vlastnosti. Princípy oboch týchto metód sú podrobnejšie popísané v [2].

##### A. Metóda DCT bez použitia pôvodného obrazu

###### 1) Proces vkladania

Pri procese vkladania sa do krycích dát I vloží vodoznak W nasledujúcim spôsobom. Krycie dáta sa najprv rozložia na bloky 8x8 a aplikuje sa na nich 2D DCT. Po tejto transformácii dostaneme maticu 512x512 koeficientov rozdelených do 64x64 blokov. Následne bol do každého bloku vložený jeden bit vodoznaku. Pre vloženie jedného bitu použijeme vždy dvojicu koeficientov a to presne koeficienty (5,2) a (4,3) resp. koeficienty 18 a 17 podľa Obr. 4.

Vkladanie prebiehalo nasledovne:

- ak bola hodnota bitu vodoznaku rovná 0 tak:
  - ak bol koeficient (5,2) < (4,3) hodnoty ostali nezmenené
  - ak bol koeficient (5,2) > (4,3) hodnoty sa navzájom vymenili
- ak bola hodnota bitu vodoznaku rovná 1
  - ak bol koeficient (5,2) > (4,3) hodnoty ostali nezmenené
  - ak bol koeficient (5,2) < (4,3) hodnoty sa navzájom vymenili

Ak však boli hodnoty oboch koeficientov rovnaké alebo veľmi hodnotami rozdielne ošetrilo sa to hodnotou "k" a to tak, že k jednému sa pripočítala hodnota "k/2" a druhému odpočítala hodnota "k/2", ale nerovnosť musela ostať neporušená. V navrhnutej metóde bolo za koeficient "k" zvolená hodnota 8 ako kompromis medzi hĺbkou vloženia a nevnímateľnosťou vloženia vodoznaku.

## 2) Proces extrakcie

Táto metóda bola navrhnutá tak, aby pri extrakcii neboli potrebné krycie dáta I ale iba označené IW a vodoznak W resp. jeho rozmer. Keďže sme pri vkladaní zabezpečili, že hodnoty koeficientov (5,2) a (4,3) sú rozdielne, preto aj proces extrakcie je pomerne jednoduchý. V prvom kroku rozdelíme označené dáta na bloky 8x8 a aplikujeme 2D DCT. Následne v každom bloku porovnávame hodnoty koeficientov nasledovne:

- ak je hodnota koeficientov (5,2)  $\geq$  (4,3) hodnota bitu vodoznaku bola 1
- ak je hodnota koeficientov (5,2)  $<$  (4,3) hodnota bitu vodoznaku bola 0

### B. Metóda SVD s použitím pôvodného obrazu

#### 1) Proces vkladania

Proces vkladania začína rozložením krycieho obrazu I a vodoznaku W pomocou SVD na tri matice U, S a V. Následne sa použijú iba matice S resp. jej singulárne hodnoty. Singulárne hodnoty krycieho obrazu modifikujú pomocou vzťahu:

$$Y_t = X_t + \alpha \cdot b_v \quad (11)$$

kde:  $Y_t$  = koeficienty po modifikácii  
 $X_t$  = koeficienty pred modifikáciou  
 $b_v$  = bit tajnej správy  
 $\alpha$  = škálovací parameter/koeficient

Takto upravená matica S sa spätne pomocou matíc U a V spojí pomocou vzťahu 1. Týmto sa zabezpečilo vloženie dát vodoznaku do krycieho obrazu

#### 2) Proces extrakcie

Proces extrakcie podobne ako vkladania začína rozložením pôvodného obrazu a označeného obrazu pomocou SVD na tri matice pričom použijeme iba matice S. Bity tajnej správy potom dosiahneme pomocou vzťahu:

$$b_v = (Y_t - X_t) / \alpha \quad (12)$$

Následne sa  $b_v$  upraví do binárneho tvaru čo zodpovedá aj binárnemu tvaru vodoznaku, ktorý bol vložený aby bolo možné následne oba vodoznaky porovnať.

## VI. VÝSLEDKY

V tejto práci sa používali obrázky lena512x512.bmp a vodoznak64x64.bmp v oboch metódach.



Obr. 5 lena512x512.bmp



Obr. 6 md64x64.bmp

A. Výsledky útoku JPEG kompresiou

DCT		
JPEG	NCC	Bitová zhoda [%]
Q85	1	100
Q70	0.8579	83.5449
Q50	0.7253	61.3281

SVD		
JPEG	NCC	Bitová zhoda [%]
Q10	0.9682	96.8018

B. Výsledky útoku otočením obrazu

DCT		
Otočenie [°]	NCC	Bitová zhoda [%]
0,1	0,9958	99,8047
0,3	0,6889	67,749
0,5	0,4998	45,9717

SVD		
Otočenie [°]	NCC	Bitová zhoda [%]
0,1	0,9943	95,9014
0,3	0,9867	98,9014
0,5	0,9814	98,7016
5	0,5389	89,4287

C. Výsledky útokov Gaussovým rozostreným obrazu bloku 5x5

DCT		
Sigma	NCC	Bitová zhoda [%]
0,5	0,9996	99,9756
1	0,9193	93,2373
2	0,5597	58,0566

SVD		
Sigma	NCC	Bitová zhoda [%]
0,5	0,9955	99,6582
1	0,9496	94,6289
2	0,9098	90,8936

VII. ZÁVER

SVD transformácia ako ukázali výsledky dokáže v určitých oblastiach vylepšiť doposiaľ používané metódy. Jednou z výhod je možnosť jej aplikácie na celý obraz poprípade na bloky rôznych veľkostí v závislosti na aplikácii. Taktiež aj vlastnosti spomenuté v III.A môžu byť využité pri návrhu algoritmu, keďže práve tieto modifikácie a úpravy používa množstvo geometrických útokov. V tabuľkách môžeme pozorovať niektoré oblasti v ktorých dosiahla SVD lepšie výsledky. Treba ale podotknúť, že prvá z metód nepotrebovala k extrakcii pôvodný obraz a druhá áno. Taktiež nevýhodou SVD je potreba matíc U a V pre extrakciu vodoznaku, takže je o niečo náročnejšia na pamäť. Existujú ale aj iné metódy vkladania práve do týchto matíc bez použitia krycieho obrazu. Na záver môžeme povedať, že je stále priestor na hľadanie nových možností pre návrh algoritmov ukrývania dát ako napríklad použitie iných doteraz menej často využívaných transformácií. Treba ale tiež povedať, že každá má svoju úlohu a oblasť v ktorej dosahuje výborné výsledky a preto je trendom využívať kombináciu rôznych typov transformácií k dosiahnutiu čo možno najlepších výsledkov.

LITERATÚRA

- [1] Foriš, P.: Perceptívne metódy digitálnej vodotlače v statických obrazoch, Dizertačná práca, KEMT FEI TU Košice, 2006.
- [2] Ridzoň.R: D Digitálna vodotlač s robustnými vodoznakmi v statických obrazoch, Dizertačná práca, KEMT FEI TU Košice, 2007.
- [3] B.Chandra Mohan, S. Srinivas Kumar : A Robust Image Watermarking Scheme using Singular Value Decomposition, JOURNAL OF MULTIMEDIA, VOL. 3, NO. 1, MAY 2008
- [4] LJIE Cao : Singular Value Decomposition Applied To Digital Image Processing
- [5] Yip Ka Wa : Singular Value Decomposition and its applications
- [6] Hovančák,R : Digitálne obrazové vodoznaky v transformovanej oblasti, Dizertačná práca, KEMT FEI TU Košice, 2004.

# Ukrývanie dát vo videosekvenciách na báze CDMA

*Gabriel BUGÁR, Vladimír BÁNOCI, Dušan LEVICKÝ, Jakub ORAVEC*

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská Republika

{gabriel.bugar, vladimir.banoci, dusan.levicky}@tuke.sk, jakub.oravec@student.tuke.sk

**Abstrakt** — Tento článok je zameraný na návrh steganografickej metódy, ktorá slúži na vkladanie tajných dát do krycích dát, ktoré sú tvorené videosekvenciou. Pri spracovaní tajných dát, ktoré sú reprezentované statickými obrazmi, sa využíva princíp CDMA. Navrhnutá metóda nahrádza dve najmenej významné bitové roviny každej snímky videosekvencie spracovanými tajnými dátami. V závere sú opísané možnosti využitia, respektíve úpravy navrhutej metódy.

**Kľúčové slová** — Steganografia, video, bitové roviny, CDMA

## I. ÚVOD

Rýchly rozvoj v oblasti informačných technológií umožňuje prenášať stále väčšie a väčšie množstvo informácií. Existuje viacero druhov týchto informácií, pričom najznámejšími sú text, statické obrazy, reč a audio. Ak prenášané dáta obsahujú viacero druhov informácií, označujú sa ako multimediálne dáta [1]. Príkladom takýchto dát je aj videosekvencia, ktorú tvorí súbor statických obrazov a prislúchajúca zvuková stopa.

Steganografia je veda, ktorá ukrýva tajné dáta do krycích dát tak, aby si to používateľ, ktorému nie sú tajné dáta určené, neuvedomil. Na rozdiel od kryptografie, kde dochádza ku šifrovaniu pôvodných dát, ktoré sa stávajú pre iných používateľov nezrozumiteľné, pri steganografii sa kladie dôraz na čo najmenšiu modifikáciu krycích dát. Tento výsledok sa dá docieľiť dvomi spôsobmi – tajné dáta sa vložia buď v časovej, alebo vo frekvenčnej oblasti.

Tento článok sa venuje metóde na vkladanie viacerých statických obrazov do nekomprimovaného videosúboru formátu AVI. Navrhnutá metóda pri spracovaní tajných dát využíva princíp CDMA, pri ich vkladaní sú modifikované všetky snímky krycej videosekvencie v časovej oblasti.

## II. STEGANOGRAFIA

Steganografia je veda, ktorá na ukrytie tajných dát využíva krycie dáta. Ako krycie dáta sa používajú textové, zvukové, obrazové, alebo video súbory. Tieto dáta sa upravujú s využitím tajnej správy tak, aby boli s pôvodnými krycími dátami čo najpodobnejšie. Využíva sa nedokonalosť ľudských zmyslov, resp. malá zmena štatistických parametrov dát. Upravené dáta sa nazývajú aj označenými.

Na steganografické systémy sú kladené viaceré požiadavky, niektoré sa vzájomne vylučujú [1].

**Nevnímateľnosť** je vlastnosť, ktorá využíva nedokonalosti ľudského zraku, resp. sluchu. Pre objektívnejšie porovnanie sa používa vyjadrená ako špičkový pomer signálu a šumu (PSNR).

**Nedetekovateľnosť** opisuje podobnosť štatistických parametrov označených dát a pôvodných krycích dát. Pri veľkých rozdieloch hodnôt štatistických parametrov sa môže zistiť použitie steganografického systému.

**Odolnosť (robustnosť)** vyjadruje mieru úsilia, ktoré je potrebné na poškodenie tajnej správy.

**Kapacita** určuje množstvo informácie, ktoré sa dá pri použití konkrétnych krycích dát ukryť.

Steganografické systémy majú viacero použití. Niektoré sú zamerané na čo najväčšiu

kapacitu systému, iné slúžia na tzv. digitálnu vodotlač. Pri nej sú systémy zamerané na čo najväčšiu odolnosť a nevnímateľnosť, resp. nedetekovateľnosť.

### III. POUŽITÉ PRINCÍPY A TECHNIKY

Táto kapitola stručne opisuje princípy a techniky, ktoré navrhnutá metóda využíva. Jedná sa o vysvetlenie parametrov videosekvencie, popísanie rozkladu na bitové roviny a vysvetlenie princípu CDMA.

#### A. Popis videosekvencie

Videosekvencia (video) sa dá opísať ako sled viacerých statických obrazov, ku ktorým môže byť pridaná zvuková stopa. Každý z týchto statických obrazov sa označuje ako snímka (anglicky frame). Medzi dôležité parametre videa patrí snímková frekvencia videa, rozmer rastra a farebná hĺbka snímok, a formát videa [2].

**Snímková frekvencia** (framerate) videa vyjadruje počet snímok videa za sekundu. Najčastejšími snímkovými frekvenciami sú frekvencie 15, 24, alebo 30 snímok za sekundu.

**Rozmer rastra snímky** udáva počet obrazových bodov snímky, je v tvare  $m \times n$ , kde  $m$  je šírka snímky a  $n$  je výška snímky. Ak sa udáva aj tretia súradnica, je to zväčša farebná hĺbka snímky.

**Farebná hĺbka snímky** hovorí o tom, koľko bitov je potrebné použiť na zaznamenanie úrovne jasovej funkcie jedného obrazového prvku snímky. Závisí od použitého farebného modelu. Pri 256 odtieňoch šedej je farebná hĺbka 8 bitov, pri modeli RGB (vyše 16,7 milióna farebných odtieňov) je farebná hĺbka 24 bitov, teda 8 bitov pre každú zložku tohto modelu.

**Formát videa** je určený kodekom, ktorý sa použije pri ukladaní videosúboru. Kodeky sa delia na dve základné skupiny – na stratové a bezstratové. Pri použití stratového kodeku dochádza ku kompresii videa, čo je vhodné kvôli zníženiu potrebnej prenosovej rýchlosti, ale z hľadiska kapacity steganografického systému to vhodné nie je.

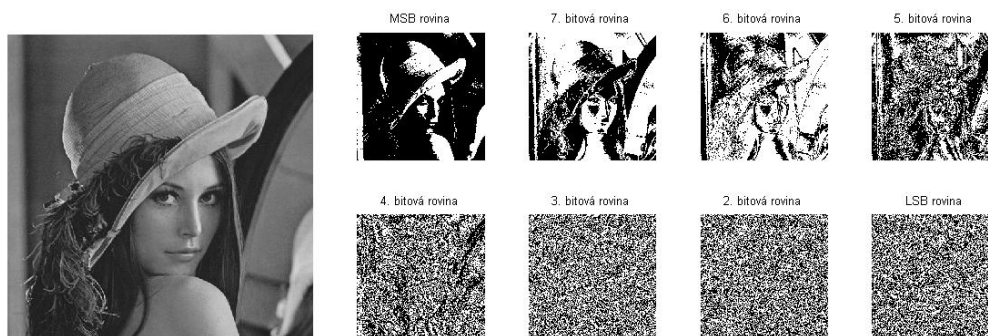
Navrhnutá metóda pracuje s videosekvenciou ako so sledom statických obrazov, ktoré ďalej spracúva. Využíva videosúbory formátu AVI, ktoré nie su komprimované.

#### B. Rozklad na bitové roviny

Pre zjednodušenie práce so statickými obrazmi sa využíva ich rozklad na bitové roviny. Počet bitových rovín statického obrazu sa rovná jeho farebnej hĺbke v bitoch. Hodnota jasovej funkcie obrazu sa pre každý obrazový prvok rozloží na mocniny dvojky. Koefficienty (0 alebo 1), ktorými sú prenasobené tieto mocniny sa uložia do obrazového prvku príslušnej bitovej roviny. V každej bitovej rovine sa tak získa tzv. binárny obraz, teda obraz, v ktorom jasová funkcia obrazových prvkov nadobúda hodnoty iba z množiny  $\{0; 1\}$ .

Je dôležité spomenúť, že jednotlivé bitové roviny nesú rôzne množstvo informácie. Sú zoradené od najdôležitejšej bitovej roviny (anglicky Most Significant Bit) po najmenej dôležitú (Least Significant Bit).

Navrhnutá metóda nahrádza dve najmenej dôležité bitové roviny snímky krycej videosekvencie tajnými dátami. To spôsobuje pri farebnom modeli RGB maximálnu možnú zmenu úrovne jasovej funkcie pre každú zložku modelu o  $3/255$ .



Obr. 1 Obraz lena.bmp a jeho rozloženie na bitové roviny



### C. CDMA (Code Division Multiple Access)

CDMA je jednou z techník, ktorá sa využíva na zdieľanie prenosového média. Používajú sa pri nej kódové slová (kódy) slúžiace na odlíšenie správ pre jednotlivých používateľov. Tieto kódy musia spĺňať podmienku **ortogonalitu**, čo v praxi znamená, že skalárny súčin ľubovoľnej dvojice kódov musí byť nulový.

Medzi ortogonálne kódy patria napr. Walsh-Hadamardove kódy. Obsahujú iba prvky z množiny {1; -1} a sú riadkami (súčasne aj stĺpcami) Hadamardových matíc. Matice sú štvorcové, dĺžky kódov sú mocniny dvojky. Prvý kód ľubovoľnej Hadamardovej matice, ktorý je tvorený samými jednotkami, sa nepoužíva, pretože nijako nemení ním rozprestierané dáta.

$$H(3) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Obr. 2 Hadamardova matica s kódmi dĺžky štyri

Pred vyslaním sa dáta pre jednotlivých používateľov spracujú. Každý bit dát sa prenášobí každým bitom kódu adresáta. Následne sa takto prenášobené dáta sčítajú. Ak máme hodnoty bitov pôvodných dát a aj prvky kódu z množiny {1; -1}, maximálna možná hodnota sčítaných dát pre  $n$  užívateľov je  $\pm n$ .

Pri prijímaní sa počíta skalárny súčet blokov prijatých dát a kódu adresáta. Získajú sa dáta, ktoré sú s pôvodnými zhodné až na amplitúdu, ktorá závisí od počtu súčasne komunikujúcich používateľov.

Navrhnutá metóda používa techniku CDMA na súčasné vloženie, resp. extrakciu troch tajných správ, ktorými sú statické obrazy.

## IV. POPIS NAVRHNUTEJ METÓDY

V tejto kapitole je opísaná navrhnutá metóda pre vkladanie dát do videosekvencií s použitím CDMA. Použitý krycí video bolo vopred nahrané webkamerou. Je dĺžky 5 sekúnd, má snímkovú frekvenciu 15 snímkov za sekundu. Rozmer rastra snímky je 320x240 obrazových prvkov, farebná hĺbka je 24 bitov (farebný model RGB). Video je uložené v nekomprimovanom AVI formáte.

Ako tajné dáta sú použité tri statické obrazy, s rozmermi rastrov 64x64 obrazových prvkov, farebnou hĺbkou 8 bitov, formátu BMP.

Metóda modifikuje dve najmenej významné bitové roviny každej snímky videosekvencie spracovanými tajnými dátami. Modifikácia každej snímky videa je síce náročnejšia na výpočtové prostriedky, ale zabraňuje jednoduchým útokom ako vystrihnutie snímky, či zmena poradia snímkov.

Na rozprestretie tajných dát slúžia tri Walsh-Hadamardove kódy dĺžky 4 (2., 3. a 4. riadok príslušnej matice).

### A. Vkladanie tajných dát

Algoritmus na vkladanie tajných dát začína ich prípravou. Každý zo statických obrazov, ktoré predstavujú tajné dáta, sa načíta a rozloží na bitové roviny. Následne sa hodnoty obrazových bodov v bitových rovinách mapujú z hodnôt {0; 1} na hodnoty {-1; 1}. Táto úprava je dôležitá z hľadiska použitia princípu CDMA. Ak by boli použité hodnoty {0; 1}, tajné dáta by neboli po prijatí správne dekódované.

Po mapovaní hodnôt sa tajné dáta preusporiadajú z matíc 64x64x8 do vektorov. Každý bit tajných dát sa prenášobí všetkými bitmi kódu adresáta (kódy sú dĺžky 4). Tým pádom sa získa vektor dĺžky  $64 \cdot 64 \cdot 8 \cdot 4 = 131\,072$  bitov. Vykoná sa súčet trojice tajných dát.

Keďže sa sčítali tajné dáta pre troch používateľov a bity jednotlivých tajných dát a aj prvky kódov boli z množiny {-1; 1}, po sčítaní sa získa jeden zo stavov {-3; -1; 1; 3}. Hodnoty jasovej funkcie obrazového prvku, ktoré budú v krycích dátach modifikované, však môžu byť len z množiny {0; 1}. Tým pádom je potrebné tieto štyri stavy zakódovať dvojicou bitov. Použitý zakódovanie stavov zobrazuje Tab. 1.

TAB. 1 KÓDOVANIE STAVOV PO SÚČTE TAJNÝCH DÁT

<b>stav po sčítaní</b>	-3	-1	1	3
<b>kód (LSB, 2.BR)</b>	0 1	0 0	1 0	1 1

Nasleduje spracovanie krycieho videa. Video sa načíta po snímkach, ktoré sú reprezentované ako matica  $320 \times 240 \times 3$ . Video sa rozloží na jednotlivé farebné zložky modelu RGB a tieto farebné zložky sa ďalej rozložia na bitové roviny. Matice rozmeru  $320 \times 240 \times 8$  sa preusporiadajú na vektory.

Pre každú bitovú rovinu snímky videa sa vytvorí vektor, ktorý obsahuje bitovú rovinu červenej, zelenej a modrej zložky snímky. Pre jednu snímku videa má tento vektor dĺžku  $320 \times 240 \times 3 = 230\,400$  bitov. Tento vektor je pre najmenej dôležitú (LSB) a druhú najmenej dôležitú bitovú rovinu nahradený vektorom spracovaných tajných dát, ktorý sa doplní nulami tak, aby bol rovnako dlhý ako vektory pre ostatné bitové roviny.

Následne sa vektor preusporiada na maticu a z jednotlivých bitových rovín sa zloží snímka videa. Každá snímka sa uloží do pamäte a po uložení poslednej sa sekvencia týchto snímkov uloží ako videosúbor.

### B. Extrakcia tajných dát

Pre extrakciu tajných dát slúži algoritmus, ktorý využíva podobné postupy ako algoritmus pre vkladanie tajných dát, ale v opačnom poradí. Po načítaní označeného videa sa video rozloží na jednotlivé snímky. Snímka sa rozdelí na zložky farebného modelu RGB a tieto zložky sa ďalej rozdelia na bitové roviny. Matice pre každú bitovú rovinu sa preusporiadajú na vektory. Z vektorov pre najmenej dôležitú a druhú najmenej dôležitú bitovú rovinu sa získa zakódovaný súčet tajných dát. Po dekodovaní sa tento súčet uloží do matice, kde sa vypočíta priemerná hodnota tohto súčtu zo všetkých snímkov videa.

Nasleduje operácia inverzná ku rozprestieraniu Walsh-Hadamardovými kódmi. Vypočíta sa skalárny súčin súčtu tajných dát a kódu adresáta. Keďže navrhnutá metóda používa Walsh-Hadamardove kódy dĺžky 4, skalárny súčin sa počíta po blokoch dĺžky 4 bity.

Po tejto operácii sa získa pre každého používateľa vektor, ktorý obsahuje bitové roviny tajných dát. Tento vektor sa preusporiada na maticu a z bitových rovín sa zostaví pôvodný obraz - tajné dáta.

## V. ZÁVER

Predložená práca sa venovala opisu metódy slúžiacej na vkladanie statických obrazov do videosekvencie s využitím princípu CDMA. Použité statické obrazy boli vo farebnej hĺbke 8 bitov a v bezstratovom formáte BMP. Spracovali sa rozložením na bitové roviny a rozprestretím Walsh-Hadamardovými kódmi. Videosekvencia obsahovala snímky s rozlíšením  $320 \times 240$  obrazových prvkov, farebnou hĺbkou 24 bitov a bola uložená v bezstratovom AVI formáte. Snímková frekvencia bola 15 snímkov za sekundu, tajné dáta boli vkladane do každej snímky videosekvencie. Takýto postup zaručí vyššiu odolnosť voči jednoduchým útokom akými sú odstránenie snímky, zmena poradia snímkov, alebo spriemernenie snímkov, ale na druhej strane podstatne zvyšuje výpočtovú náročnosť.

Navrhnutá metóda bola skúšaná aj na vkladanie tajných dát do videosekvencie v reálnom čase. Pri tomto spôsobe je dôležité minimalizovať výpočtovú náročnosť algoritmu pre vkladanie, resp. extrakciu tajných dát [3]. Navrhnutá metóda získava videosekvenciu z webkamery. Pred načítaním ďalšej snímky sa spracuje aktuálna snímka.

Pokusom bolo zistené, že navrhnutá metóda nie je vhodná pre vkladanie tajných dát v reálnom čase, keďže medzi jednotlivými načítanými snímkami dochádzalo ku veľkému oneskoreniu (približne jedna sekunda).

Navrhnutá metóda modifikuje dve najnižšie roviny každej snímky krycieho videa. Ak by bol zvolený menší počet snímkov, ktoré by sa modifikovali, výpočtová náročnosť by bola nižšia a medzi nemodifikovanými snímkami by nastalo menšie oneskorenie. Väčšie oneskorenie by však bolo stále prítomné medzi modifikovanou a nemodifikovanou snímkou, čím by sa získala nekonštantná snímková frekvencia. Taktiež by klesla odolnosť voči útokom ako zmena poradia snímkov alebo vystrihnutie snímkov.

## LITERATÚRA

- [1] LEVICKÝ, Dušan: Multimédiá a ochrana ich obsahu. Košice: Elfa, 2012. 249 s. ISBN 978-80-8086-199-5.
- [2] TOKÁR, Tamás: Digitálna vodotlač vo videosekvenciách. Dizertačná práca, KEMT FEI TU Košice, 2009.
- [3] KANÓCZ, Tomáš: Digitálne vodoznaky vo videu. Písomná práca ku dizertačnej skúške, KEMT FEI TU Košice, 2010.

# Steganografia s Využitím Retransmisie Segmentov v Protokole TCP

*Gabriel BUGÁR, Vladimír BÁNOCI, Dušan LEVICKÝ, Ondrej PALKOCI*

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská Republika

{gabriel.bugar, vladimir.banoci, dusan.levicky}@tuke.sk,  
ondrej.palkoci@student.tuke.sk

**Abstrakt** — Cieľom práce je prezentovať uplatnenie steganografických systémov v dátových sieťach. Teoretická časť opisuje charakteristické vlastnosti, metódy a základné princípy steganografie. Popisuje vytvorenie podprahového kanála v sieti a jeho použitie na steganografické účely, takzvanú sieťovú steganografiu, ktorá je pomerne mladým odvetvím a objektom výskumu v oblasti skrývania informácií. Ďalšia časť sa zaoberá spôsobmi vytvorenia skrytého steganografického kanála s využitím retransmisii v sieti. Retransmisná metóda je testovaná v praktickej časti a v závere je zhodnotený kvalitatívny prínos simulácií.

**Kľúč. slová** — Sieťová steganografia, podprahový kanál, TCP Vegas

## I. ÚVOD

História ľudstva bola neustále sprevádzaná potrebou komunikovať utajeným spôsobom. Techniky skrytej komunikácie dosahovali najväčší rozvoj počas vojen, sprisahaní, atentátov či vládných prevratov, kedy schopnosť utajenia citlivých správ znamenala rozdiel medzi životom a smrťou. Podobný charakter pretrvával po celé tisícročia, len vývoj a poznanie spoločnosti ho posúvali ďalej. Digitalizácia a príchod komunikačného veku priniesli nový pohľad na svet. Súkromie a komunikácia boli prenesené do sveta počítačov a sietí, čím sa stáva potreba ich utajovania stále aktuálnejšia. Súhrnný názov techník slúžiacich na tajnú výmenu správ je Steganografia. Táto dynamicky napredujúca vedná disciplína sa neobmedzuje len na utajenie skutočnosti, že bola odoslaná správa, ale robí prijímateľa a odosielateľa neviditeľnými. Takáto anonymita a súkromie sú v dnešnej modernej spoločnosti veľmi žiadanými. Následkom je prínos v oblasti ochrany súkromia, no na druhej strane to prináša nové možnosti ohrozenia jednotlivcov, spoločnosti a štátov. Kompromis medzi výhodami a hrozbami vytvára množstvo etických, právnych a technologických problémov.

Zámer článku je načrtnúť čitateľovi možnosti v oblasti utajovania informácií v sfére počítačov, digitálnych súborov a dátových sietí. Práca by mala priblížiť delenie, spôsoby a metódy skrytej komunikácie v digitálnych technológiách. Opíše základné všeobecné princípy steganografických systémov a ich aplikáciu do praxe. Ďalej sa bude venovať dvom hlavným odvetviam modernej steganografie, ktorými sú digitálne médiá a sieťové protokoly. Práve sieťovou steganografiou sa bude práca zaoberať podrobnejšie. Na záver by sme chceli experimentálne overiť spôsobilosť niektorého protokolu na steganografické využitie. Centrom nášho záujmu bola retransmisná steganografia. Na prenos utajených dát používa princíp retransmisii, čo je mechanizmus zabezpečenia spoľahlivého prenosu v nespoľahlivej sieti. Retransmisie sú kvôli svojmu častému výskytu v prevádzke vhodným objektom na maskovanie tajných informácií.

## II. SIEŤOVÁ STEGANOGRAFIA

Je najmladším steganografickým odvetvím, niekedy tiež nazývaným steganografia druhej generácie. Ide o skrývanie informácií s využitím sieťových protokolov ako sprostredkovateľov utajenej komunikácie. Sieťová steganografia je na vzostupe, pretože ako bolo uvedené v predošlej kapitole, digitálne covery neposkytovali dostatočnú kapacitu a bezpečnosť. Vkladanie na sieťovej úrovni prináša diametrálne rozdiely. Umožňuje prienik informácií po veľmi dlhú dobu, čo napomáha k väčšiemu prenosu dát a zároveň je možné posielat' správy v reálnom čase. Okrem toho sieťová prevádzka nie je zaznamenávaná, takže po ukončení spojenia neostávajú

žiadne stopy, ktoré by sa mohli neskôr podrobiť vyšetrovaniu. Takéto metódy sa preto veľmi ťažko odhalujú a eliminujú. Pri takejto komunikácii sa musia dodržať určité základné pravidlá, aby sa predišlo prezradeniu. Krycí protokol by mal byť sám o sebe rozšírený a v internetovej prevádzke bežný. Navyše jeho modifikácia nesmie byť abnormálna, aby pri potenciálnom sledovaní siete nevzbudila podozrenie.

#### A. Podprahový kanál

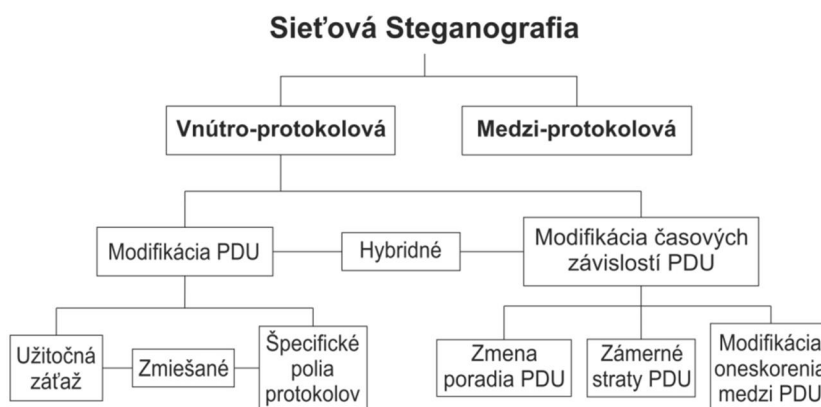
Sieťová steganografia vždy používa na tajný prenos dát podprahový, t.j. skrytý kanál. Je to mechanizmus toku informácií v rámci systému, ktorý nie je pôvodne na túto komunikáciu určený. Môže to byť akýkoľvek kanál, ktorý dokáže využiť systém na výmenu podprahových dát, bez narušenia jeho pôvodnej funkcie a účelu. Model skrytého kanála prvýkrát spomína Butler W. Lampson vo svojej práci „A Note on the Confinement Problem“ v závere roka 1973. Takisto poznamenal, že kanály je možné rozdeliť do troch kategórií a to ukladacie, legitímne a krycie. Neskôr ich v roku 2002 znova rozdelili K. Ahsan a D. Kundur [1] na ukladacie (storage channels) a časové (timing channels). Ukladací kanál je o vložení informácie do druhej pomocou zápisu do zdieľaného úložného priestoru. To znamená, že jeden subjekt vkladá dáta a iný ich následne odtiaľ číta. Subjektom v tomto prípade nemusí byť len človek, ale môže to byť aj program, alebo proces. Oproti tomu časový kanál predstavuje vkladanie informácie do druhej moduláciou dočasných systémových zdrojov. Hoci pôvodná Lampsonová koncepcia skrytého kanála nebola určená na skrytú komunikáciu, jej aplikovanie na steganografiu sa priam ponúka.

#### B. Klasifikácia

Internetová prevádzka ponúka širokú škálu protokolov a ich použitia na podprahovú komunikáciu. Jednotlivé techniky môžeme roztriediť z pohľadu modifikácie PDU (protocol data unit) [2]. Ponúka sa niekoľko možností na vytvorenie podprahového kanála a výmenu informácií. Modifikácia vlastností jedného protokolu z OSI modelu sa nazýva vnútroprotokolová (intra-protocol) steganografia. Zahŕňa napríklad zmeny v špecifických poliach PDU, vkladanie dát do užitočnej záťaže, eventuálne vykonávanie oboch týchto techník súčasne. Druhým variantom je manipulácia s časovými závislosťami medzi PDU. Tieto úpravy majú vplyv na poradie, straty alebo relatívne oneskorenia prichádzajúcich PDU. Hybridné metódy sú spojením modifikácií samotných PDU a manipulácie časových závislostí.

Využitie vzťahov medzi dvoma a viacerými odlišnými protokolmi je klasifikované ako medzi-protokolová (inter-protocol) steganografia. Pre túto metódu sa môžu použiť protokoly z rovnakej OSI vrstvy, alebo sa benefituje z väzby protokolov rôznych vrstiev. Výhodou takehoto technicky náročnejšieho prístupu, je vyššia bezpečnosť, pretože použitie viacerých protokolov znižuje schopnosť detekcie skrytého kanála. Dalo by sa povedať, že medzi-protokolová steganografia skrýva niečo, čo už bolo skryté.

Triedenie sieťovej steganografie znázorňuje nasledujúca schéma:



Obr. 1 Klasifikácia sieťovej steganografie

### III. RETRANSMISNÁ STEGANOGRAFIA

Ako to už bolo uvedené, medzi hybridné metódy patrí aj tzv. Retransmisná steganografia. Využíva princíp retransmisie v dátových sieťach za účelom vytvorenia steganografického podprahového kanála. Použitie retransmisí v dátových sieťach je veľmi obsiahla tematika budeme sa ňou zaoberať aj v ďalších kapitolách.

#### A. Princíp retransmisie

Princíp retransmisie je aparát používaný vo veľkom množstve protokolov, ktorý garantuje spoľahlivý prenos digitálnych dát, cez nespoľahlivý komunikačný kanál. Zabezpečuje ich

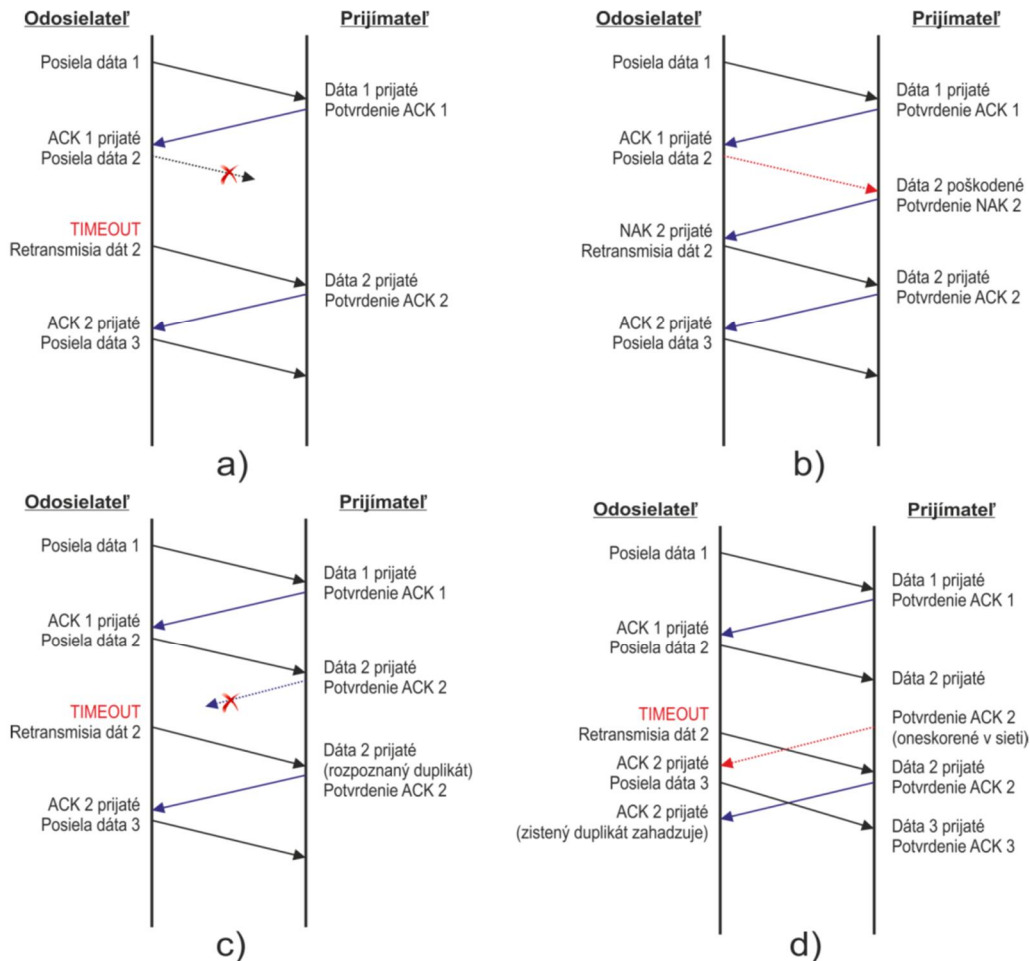
doručenie v správnom poradí, bez omeškania, poškodenia, alebo straty. Dátový prenos je totiž často ovplyvnený kanálovým šumom a počas neho môže dôjsť k degradácii, alebo zániku dát. Ak nastane podobný prípad, tak práve mechanizmy retransmisie musia zaistiť opätovné odoslanie nedoručených údajov. Spomínané mechanizmy sú dva. Prvým je posielanie potvrdzujúcich správ (acknowledgement) o úspešnom doručení na prijímacej strane. Existujú tri formy takéhoto potvrdzovania:

- Pozitívne (ACK) je spôsob, pri ktorom prijímateľ priamo potvrdzuje prijaté dáta. Nedoručenie, alebo neodoslanie ACK implicitne naznačuje stratu, alebo poškodenie odoslaných dát a potrebu ich retransmisie (napr. TCP).
- Negatívne (NAK) je posielanie správ od prijímateľa k odosielateľovi, v ktorých informuje len o chybné prijatých dátach. Často ho využívajú multicastové protokoly, pri ktorých by bolo ACK neefektívne.
- Kombinované (ACK/NAK) reprezentuje odosielanie pozitívnych, alebo negatívnych potvrdení o doručení dátach (napr. technológia Bluetooth).

Druhý mechanizmus predstavuje časovač na strane odosielateľa. Ten nastaví časový limit, počas ktorého čaká na potvrdzujúcu správu ACK od prijímateľa. Ak ju neobdrží do vypršania limitu, tak vyvolá retransmisiu. Časový limit je neustále adaptovaný na optimálnu hodnotu. Svoju úlohu pri tom zohráva RTT (round trip time), čo je doba obehu paketu v sieti. Ak by totiž bol časovač nastavený na nízku hodnotu, mohlo by to spôsobiť nadmerné množstvo zbytočných retransmisí, naopak ak by bol nastavený na veľkú hodnotu, odosielateľ by čakal prídlho na potvrdenie stratených dát, čo by znižovalo celkovú priepustnosť siete.

Sú štyri možné scenáre, po ktorých sa vyvolá retransmisia v sieti (Obr. 2):

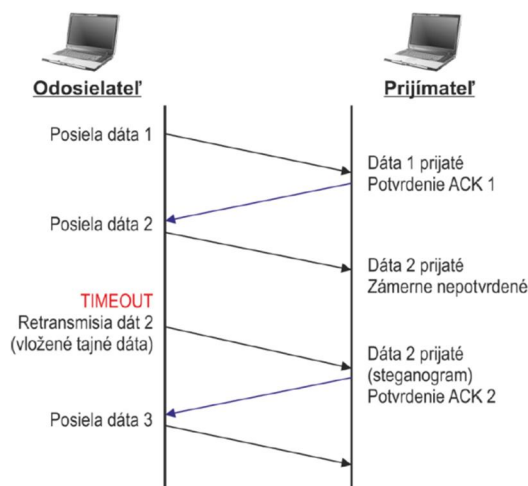
- a) Poslané dáta sa stratia počas prenosu, kvôli zahľteniu kanála. Po vypršaní limitu časovača sa dáta prepošlú.
- b) Dáta k prijímateľovi dorazia s chybou, teda počas transportu sa poškodili a prijímateľ vyvolá ich retransmisiu.
- c) Dáta dorazia k prijímateľovi v poriadku, no potvrdzujúca správa o ich prijatí sa stratí v sieti. Po vypršaní časovača sa znova prepošlú dáta na strane odosielateľa.
- d) Dáta sú doručené v poriadku a potvrdenie o prijatí je odoslané taktiež v poriadku. Zahľtenie siete však spôsobí zdržanie potvrdzujúcej správy a na strane odosielateľa vyprší limit časovača. To vyvolá retransmisiu.



Obr. 2 Scenáre retransmisí

### B. Metóda RSTEG

Touto metódou sa zaoberá viacero publikovaných článkov na internete, no ako prvý ju pod týmto názvom prezentujú Wojciech Mazurczyk a spol v [3]. Opísali metódu, kde účastníci steganografickej procedúry používajú spoľahlivý prenos dát, ktorým je posielanie súboru. V istom okamihu počas prenosu odosielateľ označí paket na retransmisiu. Modifikovaný paket na strane prijímateľa po úspešnom prijatí zámerne nepotvrdí. To by v normálnej situácii viedlo k vypršaniu časovača na strane odosielateľa a následnej retransmisii daného paketu. Avšak pri RSTEG sa pred opätovným odoslaním paketu vymení jeho užitočná záťaž za tajné dáta. Keď takto modifikovaný paket dorazí k prijímateľovi, potvrdí ho a extrahuje z neho skrytú informáciu. Tento proces je znázornený na obr. 3.



Obr. 3 RSTEG

Výkon metódy závisí od viacerých faktorov ako sú napríklad veľkosť užitočnej záťaže, rýchlosť generovania segmentov, počet vnesených retransmisií a podobne. Výber týchto parametrov musí byť optimálny a udržaný v medziach, nakoľko nevhodne zvolené hodnoty by mohli viesť k štatistickým odchýlkam a spôsobiť tak prezradenie steganografickej metódy. Vo všeobecnosti platí, že čím viac informácií skryjeme do dátového toku, tým bude jednoduchšie pre stegoanalytika odhaliť prebiehajúcu podprahovú komunikáciu. Množstvo úmyselných retransmisií nesmie presiahnuť prirodzený výskyt retransmisií v sieti, čo môže byť pri silnom zaťažení 5-7 %. Úmyselné zavádzanie je preto neustále kontrolované a dynamicky adaptované podľa aktuálnej situácie, aby sa nevybudilo podozrenie potenciálneho pozorovateľa.

### C. TCP Vegas

Vegas bol predstavený v roku 1994 ako alternatívny mechanizmus pre kontrolu zahltenia v sieti. Na rozdiel od algoritmov ostatných verzií protokolu TCP, ktoré vytvárajú zahltenie pre zistenie dostupnej kapacity, TCP Vegas predchádza vzniku zahltenia neustálym monitorovaním doby obehu paketov v prevádzke. Stratégia spočíva v nastavení rýchlosti odosielania tak, aby sa čo najmenej paketov uchovávalo vo vyrovnávacej pamäti smerovačov na prenosovej ceste. Podľa nárastu či klesania doby obehu paketu prispôsobuje veľkosť okna. Ak sa doba obehu zvyšuje, predpokladá sa zahlcovanie siete a okno sa zmenší. Naopak pokles doby obehu signalizuje možnosť zvýšiť veľkosť okna. Tento aktívny prístup protokolu TCP Vegas mu prináša zisk v podobe lepšieho výkonu a priepustnosti siete [4].

Druhú zložku algoritmu pre obmedzenie zahltenia tvorí mechanizmus pomalého štartu. Ten má za úlohu predísť poklesu výkonu TCP spojenia. Po strate paketu sa veľkosť okna zníži a postupne sa zasa zvyšuje po každom pozitívnom potvrdení zo strany prijímateľa. Oproti ostatným modifikáciám TCP je pri Vegas zvyšovanie veľkosti okna počas pomalého štartu polovičné.

Do tretice používa TCP Vegas mechanizmus, pri ktorom posudzuje kedy je efektívnejšie znova odoslať paket. Ak obdrží duplicitné potvrdenie rovnakého paketu, tak vypočíta rýchlosť jeho obehu. V prípade že prevyšuje limit časovača, TCP Vegas automaticky vyvolá jeho retransmisiu [2].

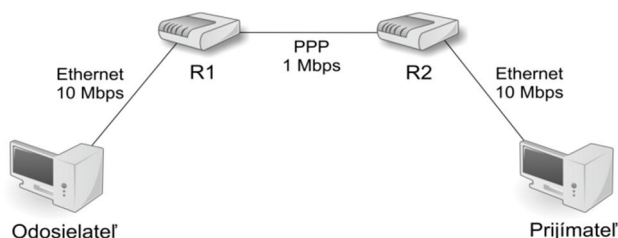
## IV. SIMULÁCIA METÓDY RSTEG

Experiment sme sa rozhodli uskutočniť v prostredí OMNeT++, čo je programové grafické rozhranie pre simuláciu komunikačných sietí. Do tohto prostredia sme implementovali knižnicu sieťových komponentov a komunikačných protokolov INET Framework vo verzií 2.2.0. Tá obsahuje oproti svojim predchodcom už aj variant protokolu TCP s prívlastkom Vegas. Práve

TCP Vegas sa stal objektom nášho experimentu. Využíva prepracovaný algoritmus obnovenia prevádzky po strate dát, preto je vhodným adeptom na steganografickú metódu.

#### A. Topológia siete

Pre lepšie porovnanie s predchádzajúcim výskumom použijeme rovnakú architektúru siete ako v [2]. Ako je vidieť na obr. 4, odosielateľ aj prijímateľ sú pripojení 10 Mbps ethernetom k prechodovým uzlom. Predstavujú ich routre R1 a R2, medzi ktorými vedie 1 Mbps PPP linka. Tá je nastavená tak, aby zavádzala oneskorenie paketov 10ms a spôsobovala 3-5 % prirodzených retransmisií. Tieto parametre sa snažia štatisticky simulovať reálne hodnoty v sieti.



Obr. 4 Scenár simulácie

Pomocou prenosu súboru bolo zahájené TCP spojenie medzi oboma účastníkmi. V konfiguračnom súbore sme nastavili MSS (Maximum Segment Size) na 1452 Bytov, čo predstavuje maximálnu veľkosť užitočnej záťaže segmentu. Taktiež sme tu zvolili použitie modifikácie TCP Vegas. Zámerné vyvolanie retransmisií pre steganografické účely, sme docielili využitím parametra PER na linke medzi R2 a prijímateľom. PER predstavuje chybovosť paketov a spôsobí ich zahodenie na prijímacej strane.

#### B. Výsledky simulácií

Danú sieť sme podrobili niekoľkým simuláciám. Prvým cieľom bolo vypočítať steganografickú šírku pásma  $S_B$  protokolu TCP Vegas. Steganografická šírka pásma  $S_B$  predstavuje prenosovú rýchlosť steganografickej komunikácie. Jej hodnotu môžeme vypočítať pomocou nasledujúceho vzorca [5].

$$S_B = \frac{N_S \cdot S_S}{T} \quad (1)$$

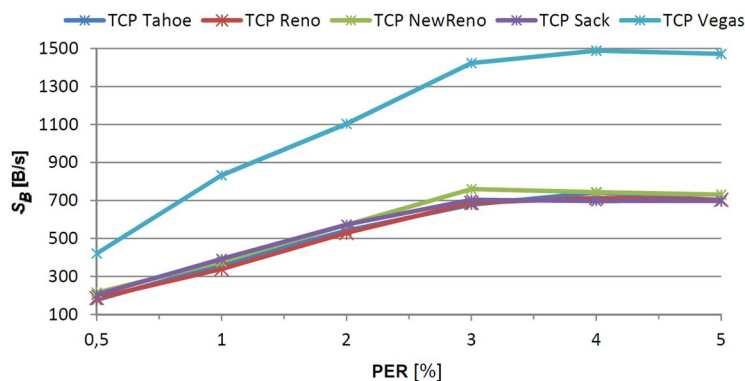
Kde  $N_S$  je počet úmyselných retransmisií,  $S_S$  predstavuje hodnotu užitočnej záťaže a  $T$  je doba, počas ktorej prebiehalo meranie.

Pri postupnom zavádzaní aditívnych retransmisií na úrovni 0.5, 1, 2, 3, 4 a 5%. Trvanie simulácie bolo stanovené na 10 minút a výsledky sú zaznamenané v tab. 1.

Tab. 1  
 Steganografická šírka pásma TCP Vegas

PER [%]	0,5	1	2	3	4	5
$S_B$ [B/s]	421,08	832,48	1103,52	1422,96	1488,61	1471,36

Z tabuľky vyplýva, že modifikácia protokolu TCP Vegas v porovnaní s inými verziami TCP, dosahuje dvojnásobne vyššiu efektívnosť ako ostatné verzie TCP. Hlavný dôvod tohto účinku, je práve spomínaný prepracovaný algoritmus obnovenia po strate paketov. TCP Vegas sa dokáže aj po úmyselne vyvolanej retransmisií pomerne rýchlo zregenerovať, čo má pozitívny vplyv na celkovú priepustnosť siete. Najväčšiu šírku pásma, až 1488,61 B/s sme získali pri zavedení úmyselných retransmisií na úrovni 4%.



Obr. 5 Porovnanie  $S_B$  protokolu TCP Vegas s ostatnými verziami protokolu TCP



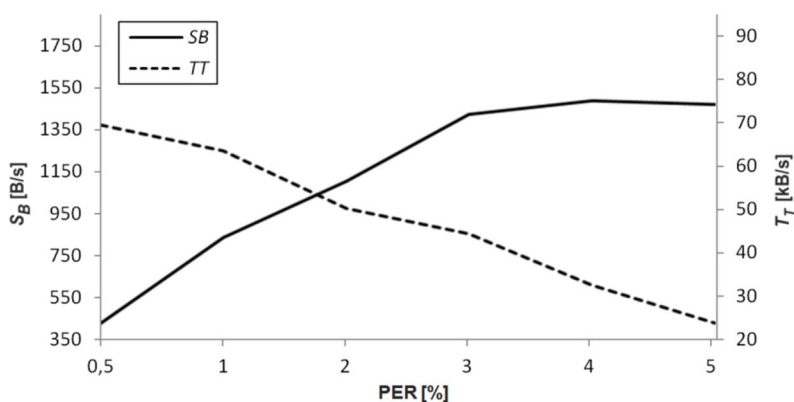
Na druhej strane treba pripomenúť, že TCP Vegas používa aj algoritmus predchádzania zahltenia. V praxi to znamená, že pri jeho použití sa do značnej miery redukuje počet prirodzených strát. Zámerné retransmisie na úrovni 4% by tak mohli vzbudiť podozrenie. Ak však obmedzíme systematické zavádzanie retransmisií na škálu 1 až 2 %, stále máme podstatne vyššiu SB, ako pri ostatných verziách TCP.

V druhej časti experimentu sme sa zamerali na vplyv prídavných retransmisií na výsledný výkon protokolu TCP Vegas. S rastúcim počtom strát paketov a ich opätovných preposielaní sa lineárne znižuje priepustnosť siete. Efektívnu priepustnosť TCP sme merali v uzle prijímateľa. Tvorí ju množstvo dát prijatých za sekundu a označuje sa ako  $T_T$  (TCP Throughput). Výsledky meraní vidíme v tab. 2.

Tab. 2  
 Vplyv retransmisií na priepustnosť TCP Vegas

PER [%]	0,5	1	2	3	4	5
$T_T$ [kB/s]	69,57	63,50	50,18	44,33	32,30	23,56

Pokles výkonu prevádzky znázorňuje aj nasledujúci graf na obr. 6.



Obr. 6 Vplyv zavádzania retransmisií na výkon TCP Vegas

Tabuľka aj graf dokazujú vyššie spomenutý argument, že zavádzanie úmyselných retransmisií zahľucuje sieť a znižuje tak jej priepustnosť. Pri hodnotách PER vyšších ako 4% klesá TT pod úroveň 30 kB/s, čo je značný pokles výkonu. Ak porovnáme priepustnosť TCP Vegas s výsledkami z [2], zistíme že dosahuje o 38-45 % vyššiu efektivitu ako najlepšie počínajúci si TCP Sack.

## V. ZÁVER

Článok sa venoval problematike aplikácií steganografických metód v počítačových a sieťových prostrediach. Práca sústreďuje svoju pozornosť na odvetvie steganografie využívajúcej pre svoje účely protokoly dátových sietí. Táto pomerne mladá oblasť utajovania informácií ponúka široké pole možností na vytvorenie podprahového komunikačného kanála.

Bližšie sme sa zaoberali aplikovaním mechanizmu retransmisie na steganografické využitie. Princíp retransmisie používa metóda RSTEG, ktorá sa radí medzi hybridné postupy sieťovej steganografie. TCP Vegas používa oproti ostatným verziám TCP sofistikovanejšie postupy, čím dosahuje lepší výkon. Algoritmus predchádzania zahltenia a schopnosť rýchlej obnovy po strate paketu v sieti, ho určujú ako vhodného adepta na retransmisnú steganografiu. Výsledky simulácií sme porovnali s predošlým výskumom a zistili sme, že TCP Vegas dosahuje v priemere až dvakrát väčšiu steganografickú šírku pásma v porovnaní s inými modifikáciami TCP. Zároveň vykazuje priepustnosť siete po zavedení úmyselných retransmisií o 38-45 % vyššiu, ako dosiahol najefektívnejší TCP Sack.

## REFERENCIE

- [1] Ahsan, K. – Kundur, D.: Practical data hiding in TCP/IP. In Proc. Workshop on Multimedia Security at ACM Multimedia, December 2002.
- [2] Jankowski, B. - Mazurczyk, W. - Szczypiorski K.: PadSteg: Introducing Inter-Protocol Steganography - In: Telecommunication Systems: Modelling, Analysis, Design and Management, Volume 58: 1-2 January/February 2015, ISSN: 1018-4864 (print version), ISSN: 1572-9451.
- [3] Mazurczyk, W. – Smolarezyk, M. – Szczypiorski, K.: Hiding Information in Retransmissions. In: Computing Research Repository.2009..
- [4] Low LP, S. - Wang, L.: Understanding TCP Vegas: Theory and practice. Submitted for publication, 2000.
- [5] Tunega, M.: Využitie retransmisie v dátových sieťach na účely steganografie. Diplomová práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2013. 63 s.
- [6] Mazurczyk, W. – Lubacz, J.: LACK – a VoIP steganographic method. Telecommunication Systems, 2010, 45.2-3: 153-163.



# Moderný Prístup Obrázovej Steganalýzy

Gabriel BUGÁR, Vladimír BÁNOCI, Dušan LEVICKÝ, Daniel MIŠENČÍK

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská Republika

{gabriel.bugar, vladimir.banoci, dusan.levicky}@tuke.sk,  
daniel.misencik@student.tuke.sk

**Abstrakt** — článok sa zaoberá steganalýzou v statických obrazoch využívajúcu transformovanú DCT oblasť, schopnú rozpoznať najpopulárnejšie steganografické algoritmy vyskytujúce sa na Internete. Začiatok práce je venovaný stručnému úvodu, základným pojmom, vlastnostiam, rozdeleniu a požiadavkám steganografického a steganalýzneho systému. Práca sa neskôr zameriava na podrobnejší opis steganalýzy v statických obrazoch na báze DCT transformácie. V ďalších častiach sa nachádza spracovanie návrhu univerzálnej steganalýzy spolu s jej vlastnosťami. V závere je možné nájsť experimentálne výsledky, ktoré zhodnocujú navrhnutú steganalýznú metódu.

**Kľúčové slová** — Steganalýza, Steganografia, Histogram, DCT transformácia

## I. ÚVOD

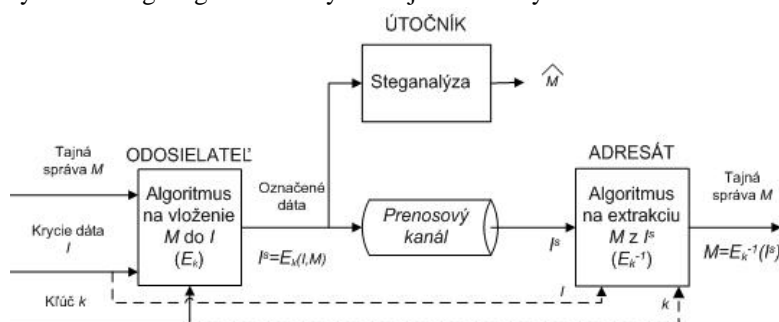
Už z histórie sú známe prvé pokusy poslať správu bezpečným spôsobom. Tak aby sa nedostala k nepovolanej osobe a aby dôležitý obsah správy nebol prezradený. Za najstaršiu techniku utajenia správy môžeme považovať kryptografiu. Kryptografia šifruje správu tak, že sa stane nezrozumiteľná. Steganografia sa snaží, aby tajná správa ostala nepovšimnutá. To je hlavný rozdiel medzi kryptografiou a steganografiou. Protikladom steganografie je steganalýza, ktorej úmyslom je odhaliť utajenú komunikáciu, teda rozličné metódy steganografie.

V posledných rokoch sme zaznamenali výrazný nárast digitalizácie nášho okolia. Nejednalo sa len o digitalizáciu textu, obrazu alebo zvuku, ale taktiež digitalizáciu našej komunikácie, vyjadrenia, presvedčenia, názoru a ich zdieľania so širokým okolím. Virtuálny priestor v dnešnom chaotickom svete ľuďom nielen uľahčuje prácu, ale ich aj zblízuje.

Na druhej strane je takáto forma virtualizácie úrodnou pôdou pre nelegálnu činnosť. Či už ide o aktivity organizujúce teroristické útoky, hackerské útoky s cieľom poškodenia alebo aktivity nabádajúce k rasovým útokom a anarchii. So zvyšujúcimi sa obavami zo strany verejnosti začali vznikať požiadavky na kontrolu tejto utajenej činnosti, ktorých výsledkom je steganalýza - veda zaoberajúca sa odhaľovaním podprahových kanálov v elektronických médiách a ich následným narušením. Steganalýza je nápomocná pri monitorovaní sieťových aktivít a tak umožňuje v predstihu zasiahnuť voči črtajúcej sa hrozbe.

## II. VŠEOBECNÝ MODEL STEGANOGRAFIE

Všeobecný model steganografického systému je zobrazený na obr. 1:



Obr. 1 Model steganografického kanála

Tajná správa  $M$  sa pomocou algoritmu na vkladanie vloží do krycích dát  $I$ . Krycie dáta  $I$ , ktoré predstavujú multimediálne dáta, sú používané a bežne dáta (text, statické obrázky, video,

reč, audio). Keďže tieto krycie dáta I sú v procese vkladania tajnej správy modifikované, je potrebné, aby obsahovali dostatočne veľkú redundanciu, ktorá sa nahradí tajnou správou M [1].

### III. STEGANALÝZA

Steganalýza je vedná disciplína, ktorej primárnou funkciou je odhalenie existencie tajnej správy v multimédiách, respektíve detekcii podprahovej komunikácie prebiehajúcej medzi dvoma stranami. Na steganalýzu je možné pozerat' ako na dvojstupňový proces: 1. Klasifikácia objektu ako stego-nositeľa, alebo nie. 2. Nájdenie pozície stego-nositeľov pixelov (t.j. pixelov obsahujúcich tajnú správu) s cieľom extrakcie, manipulácie, alebo sterilizácie tajnej správy. Ak sa v procese steganalýzy podarí odhaliť utajenú komunikáciu, považuje sa steganografický systém za prelomený a účel steganografie za zmarený [2]. Určenie obsahu tajnej správy v steganalýze predstavuje sekundárnu funkciu, ktorá vyžaduje techniky pre ďalšiu analýzu, pričom narastá výpočtová náročnosť. Všeobecne je extrakcia považovaná za doplnkovú činnosť steganalýzy. Celkovo je v steganalýze možné stretnúť sa s nasledujúcimi problémami [3]: detekcia prítomnosti tajnej správy (primárna funkcia), identifikácia algoritmu vkladania, odhad dĺžky vlozenej správy, predpoveď polohy skrytých bitov, odhad parametrov vkladania, odhad tajného kľúča použitého pri algoritme vkladania, extrakcia tajnej správy.

#### A. Všeobecná architektúra steganalýzy

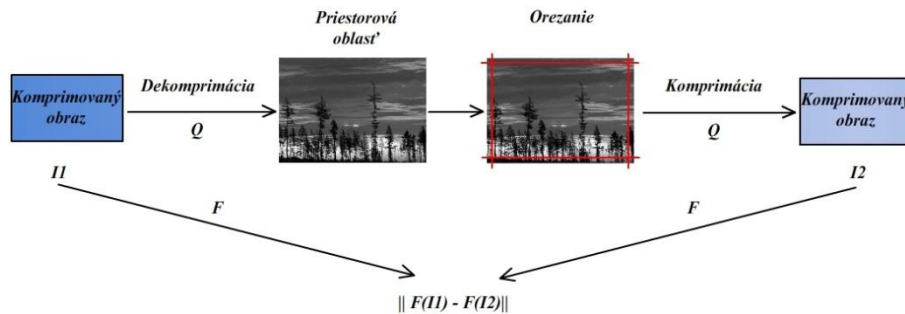
Všeobecná architektúra obrazovej univerzálnej steganalýzy zahŕňa nasledujúce hlavné kroky:

- *Predspracovanie obrazu.* Predstavuje vykonávanie operácií na poskytnutom obraze pred extrakciou príznakov, ako je konverzia RGB obrázkov do šedých obrázkov, orezanie obrázku a konštrukcia zdrojov extrakcie príznakov pomocou priestorovej a frekvenčnej transformácie. Frekvenčná transformácia získava rôzne typy podskupín vo frekvenčnej oblasti, ako napríklad predikcia zložiek koeficientov, koeficienty vysokofrekvenčných zložiek, predikcia chybných zložiek koeficientov, koeficienty zložiek obrazového šumu a log predikcia koeficientov chybných zložiek.
- *Extrakcia príznakov.* Získanie informatívnych prvkov, ktoré sú citlivé na vkladanie správy. Dobré vlastnosti by mali mať charakteristiky, ako presnosť, monotónnosť, či dôslednosť. Na základe vlastností s veľkým obmedzením, detekčný algoritmus deteguje stego obraz s veľkou pravdepodobnosťou a malou chybou. Presnosť detekcie by mala byť konzistentná s veľkým počtom setov obrazov. Funkcie by mali byť nezávislé na veľkosti obrazu, textúre, type a metóde prístupu. Vektor prvkov by mal byť monotónny pre vkladací pomer v stego obrazoch.
- *Výber a návrh klasifikátora.* Na základe extrahovaných príznakov, vybrať a navrhnúť klasifikátor, ako SVM (Support Vector Machine) klasifikátor, FLD klasifikátor, BP (Back-Propagation) neutrálny sieťový klasifikátor, geometrické klasifikátory a podobne.
- *Školenie klasifikátora.* Na školenie klasifikátora sa využívajú známe typy obrazov v tréningovom sete obrazov na tréningovanie klasifikátora a prispôsobujú sa parametre klasifikátora.
- *Klasifikácia.* Po nastavení prahovej hodnoty sa klasifikujú zistené obrázky pomocou vytrénovaného klasifikátora. Následne sa vykonáva rozhodnutie o tom, či obraz obsahuje vloženú informáciu, alebo nie. Niektoré zmiešané steganalýzne metódy vykonávajú operácie pred spracovaním na získaných príznakoch (označia ich virtuálnymi boxmi), a potom klasifikujú. „Detect images“ znamená, že získané príznaky (extrahované z obrazu, ktorý treba detegovať) sa vložia do klasifikátora pre detekciu a získava sa výstup z klasifikátora. Napríklad pre BP neutrálny klasifikátor je výstup zvyčajne desatinné číslo medzi 0 a 1. Klasifikovať znamená porovnať výstup s prahovou hodnotou a určiť, či obraz je krycí obraz, alebo stego obraz.

### IV. STEGANALÝZA ZALOŽENÁ NA BÁZE KALIBRÁCIE ŠTATISTÍK

Koncept kalibrácie bol prvýkrát publikovaný v práci [4]. Tu autori predstavili metódu pre odhad histogramu krycieho obrazu zo stego obrazu, pričom využívali steganografický algoritmus F5 [29]. Podobný princíp bol použitý aj v [5]. Tu bol úspešne napadnutý algoritmus Outguess. Kalibrácia sa preukázala ako vhodný prostriedok pre zvýšenie bezpečnosti univerzálnej steganalýzy. Jej výhodou je, že poskytuje referenčný obraz, ktorého štatistiky sú blízke štatistikám prirodzeného krycieho obrazu. Prvý univerzálny steganalýzny detektor využíval kalibráciu 23 funkcií a bol publikovaný Jessicou Fridrich v [6]. Neskôr bola táto sada funkcií rozšírená na vektor o dĺžke 274 [7]. Kalibrácia bola pôvodne zavedená pre štandard JPEG, ale je možné nájsť práce, ktoré využívajú tento koncept aj v priestorovej oblasti [8][9]. Za kalibráciu je možné považovať aj výstup z prediktora pri waveletovej steganalýze [10].

Kalibrácia v štandarde JPEG sa používa pre odhad štatistických vlastností krycieho obrazu zo stego obrazu. Tento proces je citlivejší na zmeny spôsobené vložením tajnej správy ako na zmeny v obsahu obrazu. Nevýhoda tejto metódy spočíva vo väčšej citlivosti na faktor kvality stego objektu a kombináciou rôznych faktorou kvality pri dvojito komprimovaných obrazoch. Prvým krokom pri kalibrácii jedenkrát komprimovaného obrazu je dekomprimovanie do priestorovej oblasti. Obraz je orezaný o štyri obrazové body v horizontálnom aj vertikálnom smere. Upravený obraz je naspäť komprimovaný pomocou rovnakej kvantizačnej matice do transformovanej oblasti (Obr. 18). Týmto operáciami je získaný obraz, ktorý sa svojimi štatistickými vlastnosťami veľmi podobá originálnemu kryciemu obrazu.



Obr. 2 Kalibrácie v JPEG formáte pomocou orezania obrazu

#### A. Formovanie štatistického vektora

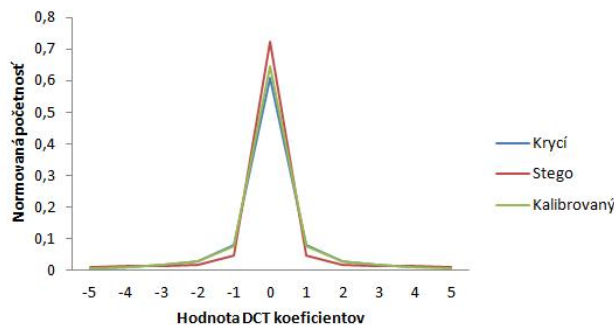
Kalibrované štatistiky sú získané ako rozdiel štatistík vypočítaných z referenčného obrazu  $I_2$  a obrazu pred kalibráciou  $I_1$  (Obr. 2).

Nech sú obrazové prvky JPEG obrazu reprezentované maticou DCT koeficientov  $d_{ij}(k)$ ,  $i, j = 1, \dots, 8, k = 1, \dots, n_B$ , kde  $d_{ij}(k)$  predstavuje  $(i, j)$  - tý kvantovaný DCT koeficient k-tého bloku, z celkového počtu blokov  $n_B$ .

Prvú sadu funkcií (11 štatistických príznakov) tvorí globálny histogram  $H$  všetkých  $64 \times n_B$  DCT koeficientov.

$$H = (H_L, \dots, H_R), \quad (1)$$

kde  $L = \min_{i,j,k} d_{ij}(k)$  a  $R = \max_{i,j,k} d_{ij}(k)$ . Na Obr. 3 je možné vidieť, že histogram kalibrovaného obrazu sa približuje k histogramu krycieho obrazu.



Obr. 3 Globálny histogram krycieho, stego a kalibrovaného obrazu

Popis výpočtu globálneho histogramu:

Premenná `DCT_koef` obsahuje všetkých 64 DCT koeficientov všetkých  $8 \times 8$  blokov obrazu.

Po výpočte histogramu (`hist`) je vybraná stredná časť daného histogramu od -5 do 5, v ktorej sa nachádza jeho maximum.

Nakoniec je stredná časť histogramu normovaná svojou početnosťou.

Ďalších päť sád funkcií ( $5 \times 11$  štatistických príznakov) predstavujú lokálne histogramy koeficientov individuálnych DCT módov  $(i, j) \in \{(1,2), (2,1), (3,1), (2,2), (1,3)\} \triangleq \zeta$

$$h^{ij} = (h_L^{ij}, \dots, h_R^{ij}). \quad (2)$$

Duálne histogramy s  $8 \times 8$  maticami  $g_{ij}^d, i, j = 1, \dots, 8, d = -5, \dots, 5$  tvoria ďalších jedenásť sád funkcií ( $11 \times 9$  štatistických príznakov):

$$g_{ij}^d = \sum_{k=1}^{n_B} \delta(d, d_{ij}(k)), \quad (3)$$

kde  $\delta(x, y) = 1$  ak  $x = y$  a  $\delta(x, y) = 0$  ak  $x \neq y$ .

Ďalšie funkcie sa získavajú z vnútro-blokovej závislosti DCT koeficientov. Prvý predstaviteľ týchto funkcií je variácia, ktorá je vyjadrená nasledovne:

$$V = \frac{\sum_{i,j=1}^8 \sum_{k=1}^{|I_r|-1} |d_{ij}(I_r(k)) - d_{ij}(I_r(k+1))| + \sum_{i,j=1}^8 \sum_{k=1}^{|I_c|-1} |d_{ij}(I_c(k)) - d_{ij}(I_c(k+1))|}{|I_r| + |I_c|}, \quad (4)$$

kde  $I_r$  a  $I_c$  označujú vektory blokov indexované  $1, \dots, n_B$  pri prehládávaní obrazu po riadkoch( $r$ ) a stĺpcoch( $c$ ). Variácia poskytuje 2 štatistické príznaky, ktoré sú v konečnom štatistickom vektore spriemerované. Výsledkom je teda jeden štatistický príznak.

Integrálne merania vnútro-blokovej závislosti (2 štatistické príznaky) sú počítané z priestorovej reprezentácie JPEG obrazu:

$$B = \frac{\sum_{i=1}^{\lfloor (M-1)/8 \rfloor} \sum_{j=1}^N |c_{8i,j} - c_{8i+1,j}|^\alpha + \sum_{j=1}^{\lfloor (N-1)/8 \rfloor} \sum_{i=1}^M |c_{i,8j} - c_{i,8j+1}|^\alpha}{N \lfloor (M-1)/8 \rfloor + M \lfloor (N-1)/8 \rfloor}, \quad (5)$$

kde  $M$  a  $N$  sú rozmery obrazu a  $C_{i,j}$  sú hodnoty jasovej zložky dekomprimovaného obrazu,  $\alpha=1,2$ .

Nasledujúca sada funkcií je vypočítaná pomocou medzi-blokovej závislosti. DCT koeficienty obrazu sú rozdelené do matice o veľkosti  $64 \times n_B$ , kde riadky tejto matice predstavujú koeficienty toho istého bloku DCT koeficientov a stĺpce tejto matice predstavujú koeficienty tej istej priestorovej frekvencie. Následne je vypočítaný rozdiel DCT koeficientov na tých istých pozíciách medzi susednými blokmi.

$$D_{i,j} = d_{i,j} - d_{i,j+1}, \quad (6)$$

kde  $D_{i,j}$  je matica, ktorá vznikne rozdielom všetkých susedných blokov DCT koeficientov pomocou horizontálneho snímania. Z tejto matice je vytvorený histogram:

$$D = (D_L, \dots, D_R), \quad (7)$$

kde  $L = \min_{i,j,k} d_{i,j}$  a  $R = \max_{i,j,k} d_{i,j}$ .

Podobne ako pri globálnom histograme, je po výpočte histogramu vybraná stredná časť daného histogramu od -5 do 5, v ktorej sa nachádza jeho maximum. Nakoniec sa využíva normovanie histogramu jeho početnosťou. Takto vznikne ďalších 11 štatistických príznakov.

Zvyšné funkcie (25 štatistických príznakov) sú počítané z matice výskytov susedov DCT koeficientov.

$$N_{00} = C_{0,0}(J_1) - C_{0,0}(J_2)$$

$$N_{01} = C_{0,1}(J_1) - C_{0,1}(J_2) + C_{1,0}(J_1) - C_{1,0}(J_2) + C_{-1,0}(J_1) - C_{-1,0}(J_2) + C_{0,-1}(J_1) - C_{0,-1}(J_2)$$

$$N_{11} = C_{1,1}(J_1) - C_{1,1}(J_2) + C_{1,-1}(J_1) - C_{1,-1}(J_2) + C_{-1,1}(J_1) - C_{-1,1}(J_2) + C_{-1,-1}(J_1) - C_{-1,-1}(J_2), \quad (8)$$

kde

$$C_{st} = \frac{\sum_{i,j=1}^8 \sum_{k=1}^{|I_r|-1} \delta(s, d_{ij}(I_r(k))) \delta(t, d_{ij}(I_r(k+1))) + \sum_{i,j=1}^8 \sum_{k=1}^{|I_c|-1} \delta(s, d_{ij}(I_c(k))) \delta(t, d_{ij}(I_c(k+1)))}{|I_r| + |I_c|} \quad (9)$$

Posledná sada funkcií predstavuje Markov model. Tento model reprezentuje rozdiel medzi absolútnymi hodnotami susediacich DCT koeficientov. Steganografické vkladanie narúša koreláciu koeficientov v krycom obraze. Poznáme dva druhy korelácie medzi JPEG koeficientmi a to interblokovú a intrablokovú.

Markov model je určený tak, že v prvom kroku sú vypočítané absolútne hodnoty z koeficientov  $F(u, v)$ . Takto vypočítané  $F(u, v)$  sú usporiadané rovnakým spôsobom ako obrazové prvky obrazu, tak že každý blok  $8 \times 8$  obrazových hodnôt je nahradený príslušnými koeficientmi. Z nich sa ďalej počítajú rozdielové polia v horizontálnom, vertikálnom, diagonálnom smere a v smere najmenšou uhlopriečkou ( $F_h(u, v)$ ,  $F_v(u, v)$ ,  $F_d(u, v)$ ,  $F_m(u, v)$ ).

$$F_h(u, v) = F(u, v) - F(u + 1, v), \quad (10)$$

$$F_v(u, v) = F(u, v) - F(u, v + 1), \quad (11)$$

$$F_d(u, v) = F(u, v) - F(u + 1, v + 1), \quad (12)$$

$$F_m(u, v) = F(u + 1, v) - F(u, v + 1). \quad (13)$$

Rozdielové polia získané zo vzorcov uvedených vyššie môžu byť modelované pomocou Markovových procesov, pričom pravdepodobnosti prechodov matice  $M_h$ ,  $M_v$ ,  $M_d$ ,  $M_m$  sú konštruované ako:

$$M_h(i, j) = \frac{\sum_{u=1}^{S_u-2} \sum_{v=1}^{S_v} \delta(F_h(u, v)=i, F_h(u+1, v)=j)}{\sum_{u=1}^{S_u-1} \sum_{v=1}^{S_v} \delta(F_h(u, v)=i)}, \quad (14)$$

$$M_v(i, j) = \frac{\sum_{u=1}^{S_u} \sum_{v=1}^{S_v-2} \delta(F_h(u, v)=i, F_h(u, v+1)=j)}{\sum_{u=1}^{S_u} \sum_{v=1}^{S_v-1} \delta(F_h(u, v)=i)}, \quad (15)$$

$$M_d(i, j) = \frac{\sum_{u=1}^{S_u-2} \sum_{v=1}^{S_v-2} \delta(F_h(u, v)=i, F_h(u+1, v+1)=j)}{\sum_{u=1}^{S_u-1} \sum_{v=1}^{S_v-1} \delta(F_h(u, v)=i)}, \quad (16)$$

$$M_m(i, j) = \frac{\sum_{u=1}^{S_u-2} \sum_{v=1}^{S_v-2} \delta(F_h(u+1, v)=i, F_h(u, v+1)=j)}{\sum_{u=1}^{S_u-1} \sum_{v=1}^{S_v-1} \delta(F_h(u, v)=i)}, \quad (17)$$

kde  $S_u$  a  $S_v$  reprezentujú dimenziu obrazu. Rozdiel absolútnych hodnôt susediacich transformačných koeficientov môže byť dosť veľký, čo by spôsobilo nárast rozmeru matíc  $M_h$ ,  $M_v$ ,  $M_d$ ,  $M_m$ . Preto autori navrhujú použiť iba centrálnu časť  $[-4,4]$  rozdielových polí. Teda ešte pred výpočtom  $M_h$ ,  $M_v$ ,  $M_d$ ,  $M_m$  sú hodnoty v rozdielových poliach  $F_h(u, v)$ ,  $F_v(u, v)$ ,  $F_d(u, v)$ ,  $F_m(u, v)$  väčšie ako 4 stanovené na 4 a hodnoty menšie ako -4 sú stanovené na -4. Takýmto spôsobom vznikne posledných 81 štatistických príznakov použitých pri detekcii. Markove modely napomáhajú zvyšovať citlivosť univerzálneho steganalyzéra pri odhaľovaní utajenej komunikácie prostredníctvom komunikačného kanála.

Spojením všetkých čiastkových funkcií vznikne finálny štatistický vektor o veľkosti 285 štatistických príznakov. Pomocou získaných príznakov klasifikátor rozhoduje, či ide o stego alebo krycí obraz.

## V. EXPERIMENTÁLNE VÝSLEDKY

Po výpočte štatistických parametrov a natrénovaní modelov je možné pristúpiť k testovaniu navrhovanej schémy. Na testovanie sa využíva databáza obrázkov, nezávislá od sady obrázkov využívaných na tréning steganalyzéra.

V tejto kapitole je porovnaný výkon navrhovaného steganalyzéra s metódou publikovanou v [10]. Úlohou je rozpoznať krycie obrazy od stego obrazov vytvorených pomocou steganografických algoritmov F5, Outguess a MHF\_DZ. Pri klasifikácii sa využíva C-SVM s radiálnym typom kernelovej funkcie. Model bol hľadaný na intervale  $\langle -3,5 \rangle$  pre parameter „C“, a  $\langle -5,9 \rangle$  pre parameter „Gama“. Vybraný bol ten, ktorý dosahoval najmenšiu cross validáciu.

Trénovacia množina obsahovala vo väčšine prípadov 148 prirodzených krycích obrazov a rovnaký počet stego obrazov, vytvorených pomocou zvoleného steganografického algoritmu s meniacou sa veľkosťou vkladanej tajnej správy. Testovacia množina obsahuje obrazy vytvorené obdobným spôsobom ako trénovacia, no s menším počtom obrazov (Tab. 1). Taktiež je zobrazený detekčný výkon systému pre dĺžku vektora 274 a 285 štatistických príznakov pri detekcii krycích a stego obrazov. Stego obrazy boli vytvorené pomocou steganografického algoritmu F5 s rôznym využitím kapacity pre vloženie tajnej správy. Vyplýva tiež, že výkon systému narastá s väčšou dĺžkou vektora. Úspešnosť detekcie stego objektov je takmer 100%. Výsledná detekcia je ovplyvnená taktiež detekciou krycích obrazov. Pri tejto detekcii bol výkon systému vyšší ak bol použitý vektor o dĺžke 285 štatistických príznakov a celkový výkon systému sa pohyboval v rozmedzí 85 až 75% v závislosti od kapacity vlozenej správy.

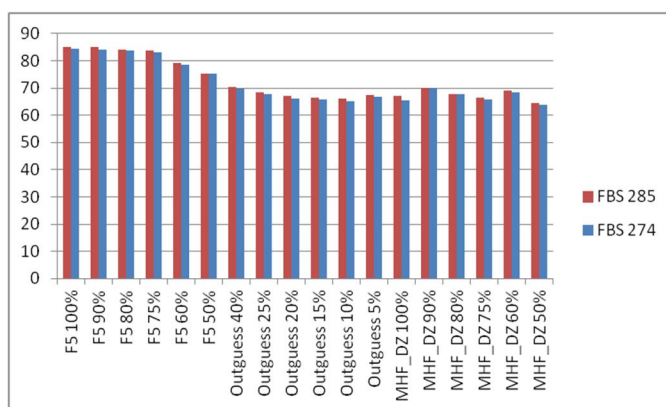
Tab. 1  
 Úspešnosť detekcie steganografického algoritmov F5, Outguess, MHF\_DZ

Kapacita [%]	F5		Outguess			MHF_DZ		
	FBS 274 [%]	FBS 285 [%]	Kapacita [%]	FBS 274 [%]	FBS 285 [%]	Kapacita [%]	FBS 274 [%]	FBS 285 [%]
100	84,24	84,92	40	69,53	70,34	100	65,46	67,14
90	84,21	84,92	25	67,75	68,52	90	69,99	70,01
80	83,67	84,20	20	66,24	66,95	80	67,87	67,78
75	83,14	83,82	15	65,63	66,31	75	65,66	66,35
60	78,50	79,18	10	65,20	65,97	60	68,48	69,16
50	75,12	75,30	5	66,59	67,36	50	63,69	64,37

Tab. 1 zobrazuje taktiež úspešnosť detekcie krycích obrazov a stego obrazov, ktoré boli vytvorené steganografickým algoritmom Outguess. Vložená správa mala menšiu veľkosť ako v prípade algoritmu F5. Z porovnania štatistických vektorov vyplýva, že vektor o dĺžke 285 štatistických príznakov dosahuje mierne lepší detekčný výkon ako vektor o dĺžke 274 príznakov. V oboch prípadoch klesá detekčný výkon so znižujúcou sa veľkosťou vlozenej tajnej

správy. Detekčný výkon systému pre stego obrázky vytvorený steganografickým algoritmom MHF\_DZ je tiež uvedený v Tab. 1. Ako v procese tréovania, tak aj v procese testovania bola využitá menšia databáza obrázkov, viď Tab. 1. Preto je možné sledovať detekčný výkon systému, nielen v závislosti od veľkosti vlozenej správy, ale aj v závislosti od počtu obrazov využitých v procese tréovania a testovania. Detekčný výkon neklesá v závislosti od veľkosti tajnej správy. Prejavuje sa tu zvyšujúci počet obrazov v databázach s menšou veľkosťou vlozenej správy. Preto sa detekčný výkon pohybuje okolo určitej hodnoty. Pre vektor o dĺžke 274 štatistických príznakov je detekčný výkon systému približne 67%. Pre vektor o dĺžke 285 štatistických príznakov je detekčný výkon systému vyšší približne na hodnotu 68%. Ďalej je možné sledovať fluktuáciu medzi detekčným výkonom jednotlivých vektorov. Najväčšie rozdiely sú pri 100% a 80% využití kapacity krycieho obrazu.

Na Obr. 4 je možné bližšie porovnať detekčný výkon navrhovanej schémy pre všetky steganografické algoritmy. Z obrázku vyplýva, že navrhovaný systém dosahuje najlepší detekčný výkon pre algoritmus F5. Vyplýva to z faktu, že pri algoritme F5 bola vložená väčšia tajná správa ako pri algoritme Outguess. Algoritmus MHF\_DZ je odolnejší voči steganalýze, lebo využíva menší počet obrazov na tréovanie systému, čo neumožňuje získanie dostatočného počtu štatistík a je modernejší ako F5.



Obr. 4 Graf klasifikácie jednotlivých druhov algoritmov

Z vyššie uvedených vyplýva, že nový štatistický vektor o dĺžke 285 štatistických príznakov prekonáva výkon štatistického vektora o dĺžke 274 štatistických príznakov. Rozdiel v detekčnom výkone je pozorovateľný pri detekcii stego obrazov vytvorených pomocou všetkých troch steganografických algoritmov využívaných v práci.

Nevýhodou väčšej dĺžky štatistického vektora je čas potrebný na výpočet štatistických parametrov a vytvorenie modelov. Pri využití vektora o dĺžke 285 štatistických príznakov bol čas potrebný na výpočet štatistických parametrov 4x dlhší, ako vektor o dĺžke 274 parametrov. Z tohto hľadiska je potrebná ďalšia optimalizácia systému pre zníženie času potrebného na výpočet potrebných príznakov. Ďalšou alternatívou pre zníženie času potrebného na výpočet parametrov je využitie iného druhu klasifikátora.

#### REFERENCES

- [1] Levický, D.: Multimédia a ochrana ich obsahu. Košice: Elfa, 2012, ISBN 978-80-8086-199-5.
- [2] Fridrich, J.: Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. In J. Fridrich, editor, Information Hiding, 6th International Workshop, volume 3200 of Lecture Notes in Computer Science, Toronto, Canada, May 23–25, 2004. Springer-Verlag, New York.
- [3] Pevny, T. – Fridrich, J.: Merging Markov and DCT features for multi-class JPEG steganalysis. In Proc. SPIE 2007 Security Steganography and Watermarking of Multimedia Contents IX, San Jose, CA, USA, Jan. 2007.
- [4] Avcibas, Ismail – KHARRAZI, Mehdi – MEMON, Nasir – SANKUR, Bülent: Image steganalysis with binary similarity measures, Uludag University, Bursa, Turkey Detection in Steganalysis, Indian Institute of Science, Bangalore, June 2006.
- [5] Prakash, G. S.: Measures for classification and detection in steganalysis, Indian Institute of Science, Bangalore, June 2006.
- [6] Farid, H. – Siwei, L. : Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines, in: F.A.P. Petitcolas (ed.): Information Hiding, 5th International Workshop. Lecture Notes in Computer Science, vol. 2578. Springer-Verlag, Berlin Heidelberg New York, 2002.
- [7] Fridrich, J. - Goljan, M. - Hoge, D.: Steganalysis of JPEG images, Breaking the F5 algorithm. In Information Hiding, 5th International Workshop, volume 2578 of Lecture Notes in Computer Science, Noordwijkerhout, The Netherlands, October 7–9, 2002. Springer-Verlag, New York.
- [8] Fridrich, J. - Goljan, M. - Hoge, D. - Soukal, D.: Quantitative steganalysis of digital images: Estimating the secret message length, ACM Multimedia Systems Journal. 9(3):288–302, 2003.
- [9] Ker, A. D.: Resampling and the detection of LSB matching in color bitmaps. In E. J. Delp and P. W. Wong, editors, Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VII, volume 5681, San Jose, CA, January 16–20, 2005.
- [10] Ker, A. D.: Steganalysis of LSB matching in grayscale images, IEEE Signal Processing Letters. 12(6):441–444.

# Position Controller for Permanent Magnet DC Machine

Karol KYSLAN, František ĎUROVSKÝ

Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

karol.kyslan@tuke.sk, frantisek.durovsky@tuke.sk

**Abstract** — The paper addresses design of position controller for the permanent magnet DC machine. Combination of P-controller for position and state controller for speed is used and calculation of its parameters is presented. In addition, values of the feedforward signals were designed in order to allow precise position tracking of S-curve motion profile. Proposed position controller was validated by simulation in Matlab/Simulink.

**Keywords** — controller, feedforward, permanent magnet DC machine, position control

## I. INTRODUCTION

Despite the fact that AC machines now covers the major part of machine market, DC machines, especially permanent magnet DC machines (PMDC) still have had their field of use in various types of applications, beginning with the machine tools and small home appliances, continuing to automotive solutions, robotic applications, photovoltaic applications, etc. [1]. Linear speed-torque curve, excellent overload features, high torque at low speed, simplicity of torque/speed control, are very known advantages of the PMDC machine [2]. Mathematical model of PMDC machine is easy to understand and though these machines have been extensively studied and many advanced control strategies have been presented in last few years. In this paper, a classical position control of PMDC machine is presented with the impact on controller's parameters design and calculation [3]-[6]. The main goal is to design a controller which would be able to track S-curve position profile.

## II. MATHEMATICAL MODEL OF PMDC MACHINE

Mathematical model of PMDC is based on the following assumptions:

- sample and hold elements of the A/D and D/A converters are neglected,
- dynamics of discrete filter for current and speed filtering are neglected,
- H-bridge converter has fast dynamics and so its time constant can be neglected,
- parameters of the machine do not vary with time,
- flux produced by permanent magnets is constant,
- nonlinear friction torques (static, slip-slick etc.) are added to the load torque,
- voltage drop across transistors and brushes is neglected,
- current of converter is unlimited (i.e. no saturation block).

Under these assumptions mathematical model of PMDC is assumed to be linear and stated as follows:

$$v_a = Ri + L \frac{di}{dt} + K_e \omega \quad (1)$$

$$J \frac{d\omega}{dt} = K_t i - B\omega - T_L \quad (2)$$

$$\frac{d\varphi}{dt} = \omega \quad (3)$$

TABLE 1. PARAMETERS OF THE PMDC MACHINE

$v_a$	applied voltage	[V]
$i$	armature current	[A]
$R$	armature resistance	[ $\Omega$ ]
$L$	armature inductance	[H]
$K_e$	back-emf constant	[Vs/rad]
$K_t$	torque constant	[Nm/A]
$\omega$	angular speed	[rad/s]
$\varphi$	angular position	[rad]
$B$	viscous friction coefficient	[Nms/rad]
$T_L$	applied load torque	[Nm]
$J$	total inertia mass	[kg/m <sup>2</sup> ]

### III. DESIGN OF CONTROLLERS

#### A. Design of current controller

Current controller was designed by applying inverse dynamics method. PMDC machine was supplied with H-bridge converter in order to ensure four quadrant operations. Converter was controlled according to bipolar control strategy and discrete PI controller was used for the current control. When omitting the converter dynamics, and assuming only gain of the converter  $K_C$ , scheme for current control synthesis is depicted in Fig.1. Back-emf voltage will be later compensated, therefore we assume neglected back-emf voltage  $v_{emf}$  and open loop transfer function (OLTF) of the current is:

$$G_I(s) = \frac{K_{rl} K_C K_a (1 + sT_{rl})}{sT_{rl} (1 + sT_a)} \quad (4)$$

we put  $T_{rl} = T_a$  and than closed loop transfer function (CLTF) of the current loop becomes:

$$\frac{i(s)}{i_{ref}(s)} = \frac{1}{1 + sT_n}; \quad T_n = \frac{T_{rl}}{K_{rl} K_C K_a} \quad (5)$$

where  $T_n$  is desired time constant of the CLTF. Thus, parameters of PI controller are:

$$T_{rl} = T_a; \quad K_{rl} = \frac{T_{rl}}{K_a K_C T_n} \quad (6)$$

Influence of the back-emf voltage on the current controller is cancelled with feedforward compensation voltage  $v_{c1}$ , application of this feedforward voltage is obvious from Fig.1. Note that compensation of back-emf voltage works only when the signal of actual speed  $\omega$  is available.

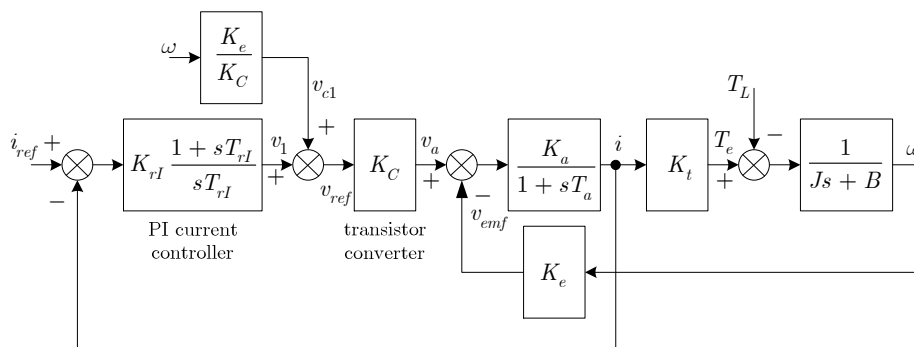


Fig. 1. Current controller design scheme



### B. Design of position controller

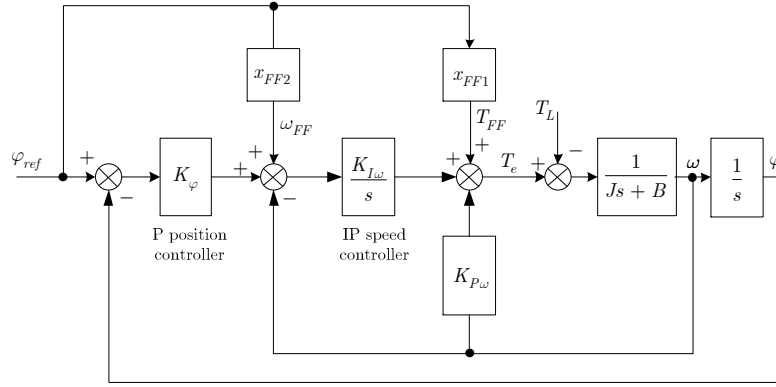


Fig. 2. Position controller design scheme

Let's assume very small value of desired time constant of current CLTF ( $T_n \rightarrow 0$ ) and output of position+speed controller is reference torque  $T_{ref}$ . Than, the reference and actual torque is assumed to be equal ( $T_{ref} = T_e$ ) because of small  $T_n$  and  $K_t$  being cancelled before and after current CLTF. For the position control of PMDC machine, a structure according to [4] is used, depicted in Fig. 2. This structure uses combination of P-controller of position and state controller (in [4] called also as IP controller) of speed and has very useful behaviour. CLTF of actual position with unknown feedforward speed ( $\omega_{FF}$ ) and torque ( $T_{FF}$ ) signals from reference position is ( $x_{FF1}$  and  $x_{FF2}$  are unknown values of feedforward):

$$\frac{\varphi(s)}{\varphi_{ref}(s)} = \frac{\frac{x_{FF1}}{J}s + \frac{K_{I\omega}x_{FF2}}{J} + \frac{K_{I\omega}K_\varphi}{J}}{s^3 + \left(\frac{B + K_{P\omega}}{J}\right)s^2 + \frac{K_{I\omega}}{J}s + \frac{K_{I\omega}K_\varphi}{J}} \quad (7)$$

we choose  $x_{FF2} = s$  and then get:

$$\frac{\varphi(s)}{\varphi_{ref}(s)} = \frac{\left(\frac{K_{I\omega} + x_{FF1}}{J}\right)s + \frac{K_{I\omega}K_\varphi}{J}}{s^3 + \left(\frac{B + K_{P\omega}}{J}\right)s^2 + \frac{K_{I\omega}}{J}s + \frac{K_{I\omega}K_\varphi}{J}} \quad (8)$$

it can be easily shown, that choosing:

$$x_{FF1} = Js^2 + (B + K_{P\omega})s \quad (9)$$

ensures desirable position tracking (i.e. unity gain of position CLTF). Feedforward signals were calculated with respect to the reference position, and thus equations for feedforward signals are:

$$T_{FF} = J\varepsilon_{ref} + (B + K_{P\omega})\omega_{ref}; \quad \omega_{FF} = \omega_{ref} \quad (10)$$

Choosing desired pole location as follows:

$$k_{1,2} = -d\omega_0 \pm j\omega_0\sqrt{1-d^2}; \quad k_3 = -k\omega_0 \quad (11)$$

desired polynomial is defined:

$$P(s) = (s - k_1)(s - k_2)(s - k_3) = s^3 + [\omega_0(2d + k)]s^2 + [\omega_0^2(2dk + 1)]s + k\omega_0^3 \quad (12)$$

where  $d$  is desired damping coefficient,  $\omega_0$  is natural frequency and  $k$  is a movement of the 3<sup>rd</sup> pole. Now values of controllers coefficients are obtained by comparing coefficients in desired and true polynomial:

$$K_{P\omega} = J\omega_0(2d + k) - B; \quad K_{I\omega} = J\omega_0^2(2dk + 1); \quad K_\varphi = \frac{k\omega_0}{2dk + 1} \quad (13)$$

It can be seen in (13) that the coefficients of the position controllers depend only on two mechanical parameters  $J, B$  and three optional parameters  $d, \omega_0$  and  $k$ .

#### IV. SIMULATION RESULTS

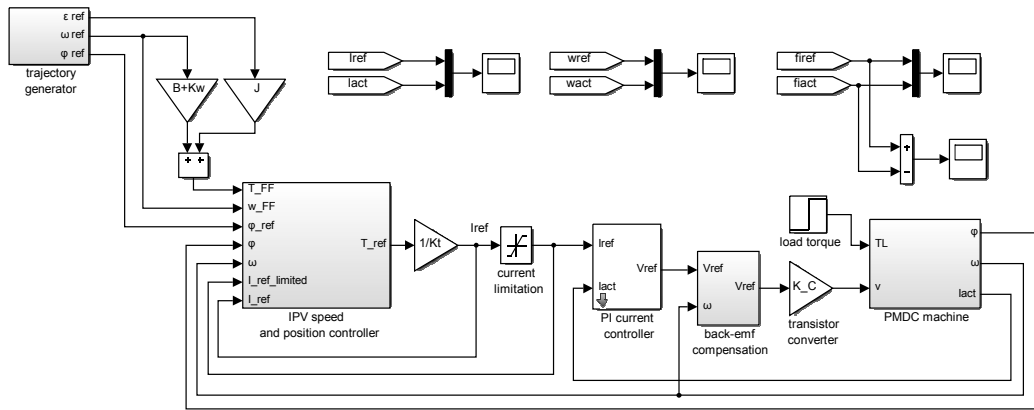


Fig. 3. Simulation scheme of PMDC position control

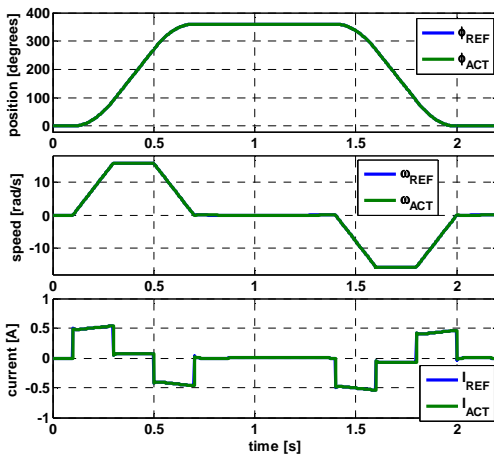


Fig. 4. Position, speed and current responses under no load parameters:  $d=0,95$ ;  $\omega_0=20$  and  $k=5$

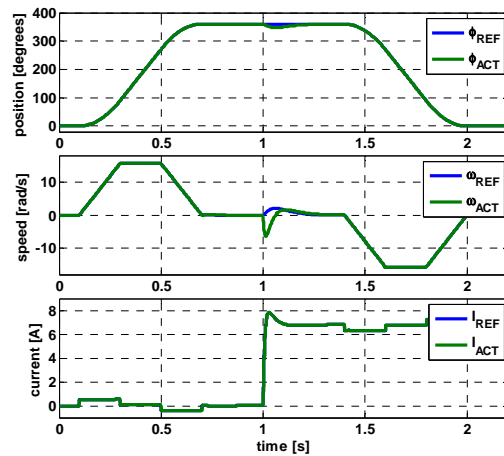


Fig. 5. Position, speed and current responses under loading parameters:  $d=0,95$ ;  $\omega_0=20$  and  $k=5$

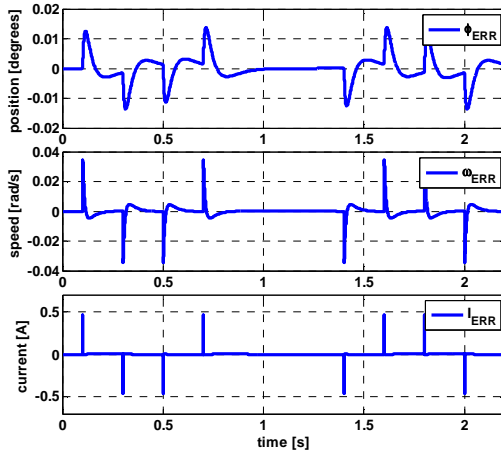


Fig. 6. Position, speed and current error under no load parameters:  $d=0,95$ ;  $\omega_0=20$  and  $k=5$

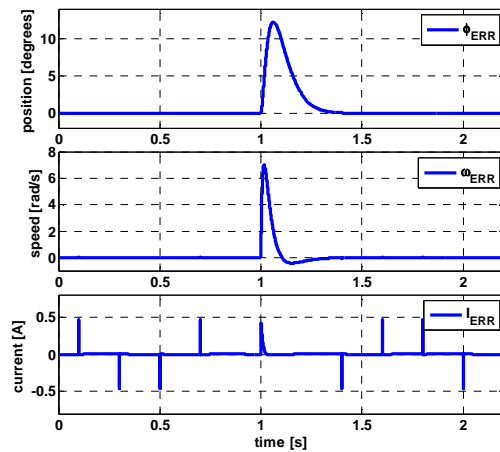


Fig. 7. Position, speed, and current error under loading parameters:  $d=0,95$ ;  $\omega_0=20$  and  $k=5$

Proposed position controller was simulated in Matlab/Simulink according to simulation scheme in Fig. 3. In all simulation cases the shaft was turned one revolution clockwise and then back. *Trajectory generator* represents block which generates so-called motion profiles. These are reference values of angular acceleration  $\epsilon_{ref}$ , from which reference speed  $\omega_{ref}$  and reference position  $\phi_{ref}$  are obtained by integration. It results in widely-used S-curve motion profile for reference position.

Actual values of particular variables when no load torque (no disturbance) is acting are depicted in Fig. 4 and their errors, as a difference between reference and actual values, are depicted in Fig. 6. In Fig. 5 and Fig. 7 is the same situation, but load torque  $T_L = 0,5$  Nm was applied on

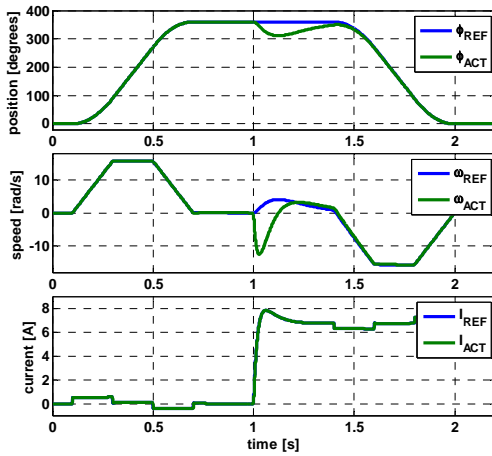


Fig. 8. Position, speed and current responses under loading, parameters:  $d=0,95$ ;  $\omega_0=10$  and  $k=5$

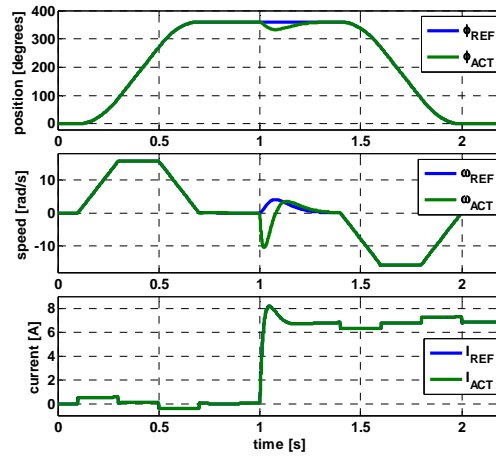


Fig. 9. Position, speed and current responses under loading, parameters:  $d=0,95$ ;  $\omega_0=20$  and  $k=2$

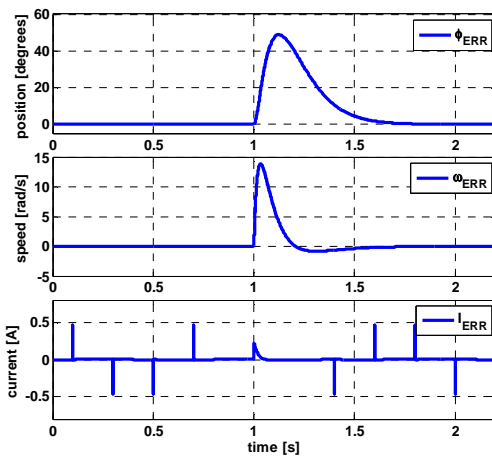


Fig. 10. Position, speed and current error under loading, parameters:  $d=0,95$ ;  $\omega_0=10$  and  $k=5$

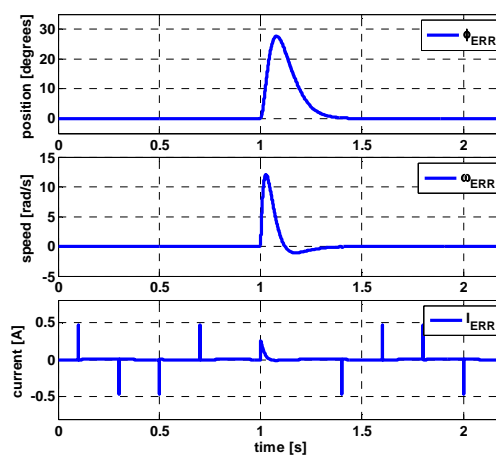


Fig. 11. Position, speed and current responses under loading, parameters:  $d=0,95$ ;  $\omega_0=20$  and  $k=2$

shaft. It can be observed, that the largest error value of position in Fig. 6 is 54 arsec, but in Fig. 8 is even up to 12 degrees.

An influence of change of parameters  $\omega_0$  and  $k$  were observed in Fig. 8 to Fig. 11. In Fig. 8, a natural frequency was changed from  $\omega_0=20$  to  $\omega_0=10$  (compared to Fig. 5). Note, that position error is considerably increased, what is depicted in Fig. 10. In Fig. 9, a movement of the 3<sup>rd</sup> pole, parameter  $k$  was changed from  $k=5$  to  $k=2$  (compared to Fig. 5). Again, position error is increased, what is depicted in Fig. 11.

### CONCLUSION

It is obvious, that when parameter  $k$  is increasing (e.g. 3<sup>rd</sup> pole is moving to the left on x-axis), absolute position error is decreasing and the dynamics of the error rejection stays approximately the same. When parameter  $\omega_0$  is increasing, absolute position error is decreasing, but dynamics of the error rejection changes. These conclusions are based on multiple simulations with many values of parameters  $k$  and  $\omega_0$ . The precise determination of the stability of proposed controller based on root locus is beyond the scope of this paper. Furthermore, tuning of the feedforward signals is possible, but it has not been considered here.

It should be appreciated, that proposed position controller is ideal, with many simplifications assumed. In real application, many of the assumptions from Chapter II are no longer valid and thus practical implementation needs more precise analysis. Benefit of the proposed controller is that it has a smaller number of coefficients than pure state-space controller and yet it preserves the dynamics of the state space controller.

#### ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract no APVV-0185-10. This work was also supported by Slovak Cultural and Educational Agency of the Ministry of Education of Slovak Republic under the contract KEGA 042TUKE-4/2012 "Teaching innovation in control of mechatronic systems".

#### REFERENCES

- [1] Klein, Frank N.; Kenyon, Mark E., "Permanent Magnet DC Motors Design Criteria and Operation Advantages," *IEEE Transactions on Industry Applications*, vol.IA-20, no.6, pp.1525, 1531, Nov. 1984.
- [2] Dúbravský, J; Príspevok k metódam a algoritmom riadenia jednosmerných servopohonov. *Dizertačná práca*. STU v Bratislave, 2012.
- [3] Žalman, M.; *Akčné členy*, STU v Bratislave, 2002.
- [4] Žalman, M.; *Riadenie pohybu (3)*. ATP Journal, 4/2009.
- [5] Malek, M., Makyš, P., Štulrajter, M; „Feedforward control of electrical drives - Rules and limits“. *Advances in Electrical and Electronic Engineering*, 9 (1), pp. 35-42., 2011.
- [6] Bocker, J.; Beineke, S.; Bahr, A., "On the control bandwidth of servo drives," *13th European Conference on Power Electronics and Applications, 2009. EPE '09*, pp.1-10, 8-10 Sept. 2009.

# Utilisation of network simulators for teaching computer network oriented subjects

*Marek POLOMSKÝ, Miroslav MICHALKO*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

marek.polomsky@student.tuke.sk, miroslav.michalko@tuke.sk

**Abstract** — this paper analyse functionality of several network simulators available at the market and compares its usability as supporting material for training in computer networks oriented subjects. The paper outlines cons and pros of those simulators in brief. Paper also describes tasks creation in Cisco Packet Tracer tool which was selected as supporting tool for student testing during classes and self-testing for distance learning.

**Keywords** — Network simulator, Cisco Packet Tracer, GNS3, NS-2, NS-3, Opnet Modeler

## I. INTRODUCTION

Nowadays, there are numerous simulators suitable for an usage in computer networks. Network simulators play significant role in proposal and testing of new network technologies, protocols and standards, what helped forming the Internet of today. Simulators can be used to design new network topologies, adjustment and testing of current ones or even adjustment and testing of its configurations. In addition to a private usage, simulators are suitable for educational purposes as well, hence they can be used as replacements of very expensive network devices which are unavailable in many cases.

## II. NETWORK SIMULATORS

Simulation is generally used in many fields of science and technology industry in order to save time and resources. During simulation, computer models not only hypothetical situations, but the real ones as well and observation of various parameters leads into assumption of the final behaviour. Network simulator is a very handy tool allowing us to simulate behaviour of various network topologies and protocols at home or in the work, excluding the need of expensive network devices or even modification of existing working network topologies in order to redesign and test it. Requirements to run such simulator consists of a software for particular simulator and a computer equipped with adequate software and hardware resources. Chosen simulator are briefly described below.

### A. Nsnam network simulators

Freely available and very popular network simulators containing libraries with plenty of network components.

#### NS-2

NS-2 is a continual object-oriented network simulator. The base of this simulator is implemented in object-oriented language C++ and scripts are executed by OTcl interpreter. Usage of such programming languages is very important because of their internal properties. The simulator runs by default on Linux but it is executable on Windows as well when emulated by emulators. Events scheduler and simple network components are implemented in C++ what results in a reduction of packets and truncation of its process time. That is the reason why C++ language is used during execution of detailed protocols and OTcl is used for a user control of simulation. It is particularly suitable for simulation of local and wide area networks [7].

Opensource property ensures fast development and frequent updates thanks to thousands of volunteers and non-profit organizations. This particular simulator helped to form the Internet of

nowadays because it was used by IETF organization to simulate many network standards prior to release. NS-2's library may be statically or dynamically connected to a main program in C++. This main program defines topology of simulation and runs simulator [1].

### NS-3

NS-3 was created as a successor of NS-2 because of its construction limitations. Simulator NS-3 is a new simulator, it is faster and more flexible at the expense of incompatibility with NS-2. Its design modifies a lot of mechanisms based on acquired experiences from NS-2. Documentation is more complex and complete as well. The core of NS-3 is programmed in languages C++ and Python, where in the case of C++, some of design patterns were used. Topologies are created by scripts using mostly prepared modules and libraries. These scripts are consecutive run under the Waf framework [6].

NS-3 is an active project, still in development, developed and used by a number of volunteers, developers and other users giving feedback [8].

### B. Cisco Packet Tracer

This simulator consists of Cisco devices only. It is freely available for members of Cisco Network Academy for educational purposes. It is implemented for Linux and Microsoft Windows. Besides graphic user interface it is interesting for its functionality as well. Devices are scalable hence the user is given an opportunity to use devices that are really available on the market. For example, use of expanding modules results in a reduction of a number of devices in topology, what would in real world result in a reduction of expenses [2].

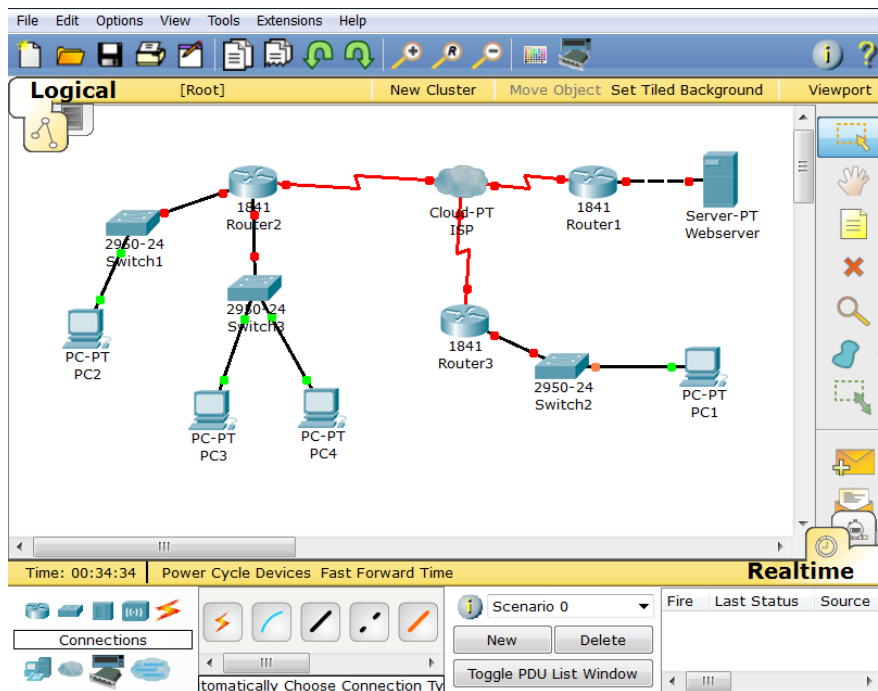


Fig. 1 Cisco Packet Tracer user interface

### C. GNS3

It is graphic network simulator freely available under GPL licence. It is capable of simulating graphic models of networks of various complexity by virtual interface running on a computer. This interface emulates network devices which run the real operating system IOS. The option of using all of the IOS commands is an advantage resulting in real device IOS feedback. Emulation is realized by main program Dynamips, what is capable of emulating IOS on a standard computer. There are other emulation applications like QEMU, Pemu, VirtualBox, which can emulate devices like Cisco ASA and, PIX firewalls, router and other network devices by company Juniper [3][4].

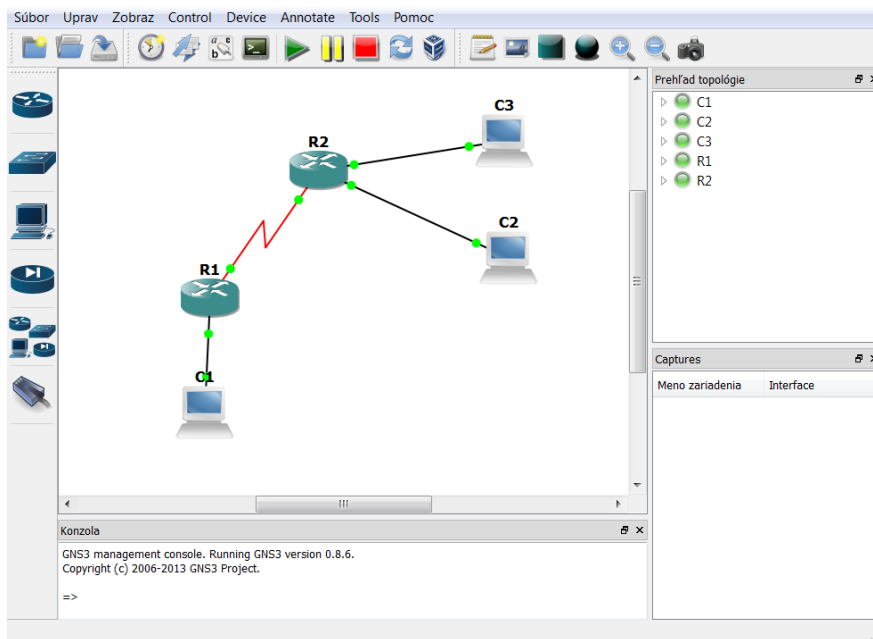


Fig. 2 Preview of user interface in GNS3

#### D. Opnet Modeler

Opnet Modeler Suite is a simulation software by Riverbed available by licencing. For seek of educational purposes, IT Guru Academic Edition is available offering 6 months free renewable licence. Opnet Modeler is capable of analysis of real simulation networks with consecutive comparison of various technology design impacts. Library consists of numerous models of devices and various brands including Cisco devices.

The simulation tool can test and demonstrate various technology designs in topologies while simulating real conditions prior to installation. Results of simulation may be displayed in a form of graphs or various statistics. This may lead in an acceleration of network topology designing. That is one of the reasons why it is one of the most popular network simulators [1][5].

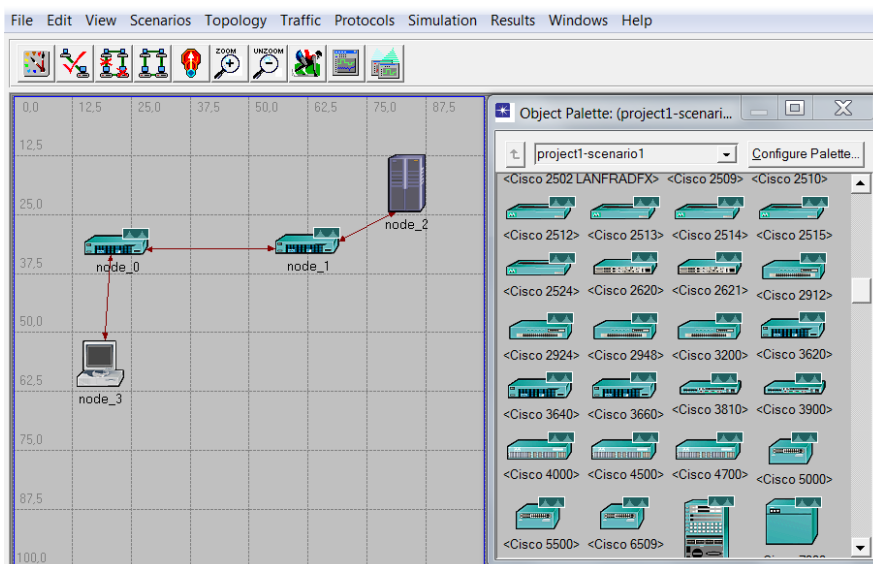


Fig. 3 Opnet Modeler user interface via IT Guru Academic Edition

### III. SELECTING THE BEST SOLUTION FOR DCI NETWORK SUBJECTS

Numerous parameters have to be emphasized when selecting the best simulator for creation of materials according to the needs of school subjects Computer networks and Application of computer networks. These parameters are price, availability, functionality, orientation, graphic user interface and its manageability and options.

NS simulators are, according to our needs, unusable because interface of NS simulators is controlled by complicated commands. NS3 does not support Cisco devices which are closely connected with Computer networks and Application of computer networks subjects.

Simulator Opnet Modeler is a professional tool but a very complicated tool as well due to its complexity. Its setup and control are very complicated and tedious. Its IT Guru Academy licence is associated with each student and have to be updated every 6 months. IT Guru Academy licence is limited by maximal import and export size, maximal crossing nodes of 20 and not supporting other modules. Another negative aspect is interrupted support of IT Guru Academy what means no future updates. Purchase of licences would be very financially demanding and its potential would remain not fully exploited.

GNS3 is freely available and its installation is easy. This simulator need to be setup prior to the start. Initialisation is followed by a wizard. The first step is a setup of IOS image path and consecutive execution what starts an optimization of devices assigned with particular IOS image. Simulator may be used afterwards. System demands rise by rising number of devices in topology, so it is advised to run optimization eventually. Its advantage lies in an option, allowing to configure network components directly on the emulated IOS system what results in a functionality of all CLI commands supported in the particular IOS image.

In contrast with GNS3 simulator, Cisco Packet Tracer is available just after the registration in Cisco Netacad academy. Its advantages are little system requirements and ease of installation. Application is ready to run after the installation. Devices just being used in the topology are displayed in life-size in an interactive image and capable of being hardware updated if possible. As the only one of all the simulators, it offers an option of settings check according to prepared assignment. It is not possible to process some of the commands because of limitations and the fact, that the simulator is developed for the Cisco academy purposes only, what is a disadvantage.

I decided to create materials for Cisco Packet Tracer and GNS3 according to my experiences with simulators. These simulators do have the most advantages out of described ones. Their installation and consecutive control is very simple and their GUI is clear and dynamic. Configuration of network devices is realised directly by CLI commands which are used in the same way in real Cisco devices.

#### IV. CREATION OF TASKS FOR STUDENTS

As Packet Tracer was selected more than twenty tasks were created for students to subject Computer Networks and Application of Computer Networks. Fig. 4 show testing topology for students to test Frame Relay understanding.

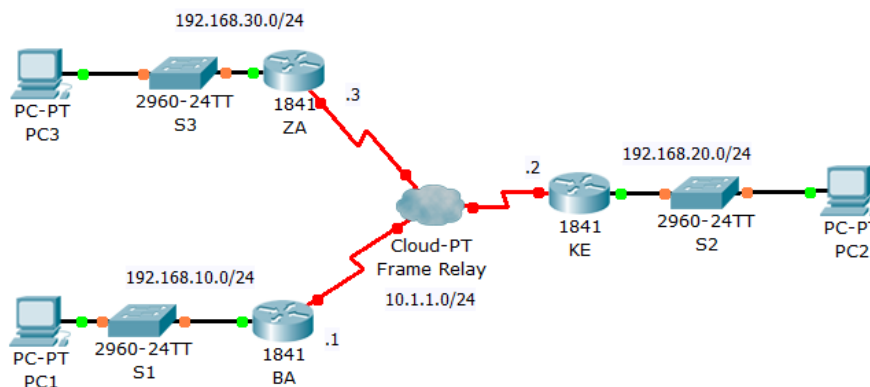


Fig. 4 Frame Relay sample testing topology created in Cisco Packet Tracer

One of the key feature of Packet Tracer is evaluation and grading system using pre-defined scoring system. Student is scored in real-time or after submission based on configures behavior. Fig. 5 shows user interface for grading point's definition. Resulting tasks are published to students using official website of subjects Computer Networks and Application of Computer Networks. Students are free to make self-testing in order to receive valuable feedback of their knowledge understanding. Some Packet Tracer activities were created as final exams at the closing subject evaluation.





Fig. 5 Grading options for Activity Wizard

## V. CONCLUSION

Several simulators available on the market nowadays were described in this work. Based on their advantages, disadvantages and consecutive comparison, it was decided, which of the described simulators are the most suitable for the creation of materials for the selected school subjects. The most suitable solution is simulators Cisco Packet Tracer and GNS3 because of their simple use and advantages over the other ones.

## ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICALOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is financed from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] ALTMAN, Eitan, JIMÉNEZ, Tania 2012. Simulator for Beginners. San Rafael: Morgan 2012, ISBN: 978-160-8456-925
- [2] Cisco Packet Tracer Dokument [online]. Available online: <[https://www.netacad.com/documents/300010/11300388/Cisco\\_PacketTracer\\_AAG.pdf/39c04df5-0cf0-4267-81b8-9cd98c1f43f8](https://www.netacad.com/documents/300010/11300388/Cisco_PacketTracer_AAG.pdf/39c04df5-0cf0-4267-81b8-9cd98c1f43f8)>.
- [3] GNS3 Documentation, [online] [cit. 15.12.2013]. Available online: <[http://icourse.cuc.edu.cn/computernetworks/cisco/Simulator/Dynamips/GNS3-0.4\\_documentation.pdf](http://icourse.cuc.edu.cn/computernetworks/cisco/Simulator/Dynamips/GNS3-0.4_documentation.pdf)>.
- [4] GNS3 Graphical Network Simulator, Introduction to GNS3 [online] [cit. 17.11.2013] Available online: <<http://www.gns3.net/documentation/gns3/introduction-to-gns3/>>.
- [5] LU, Zheng, YANG, Hongji 2012. Unlocking the Power of OPNET Modeler. New York: Cambridge University Press, 2012. ISBN 978-0-521-19874-
- [6] NS by example [online] [cit. 25.11.2013]<<http://nile.wpi.edu/NS/>>
- [7] The Network Simulator NS-2, official website [online] [cit. 21.11.2013]. Available online: <<http://www.isi.edu/nsnam/ns/>>.
- [8] The Network Simulator NS-3, manual [online] [cit. 21.11.2013]. Available online: <<http://www.nsnam.org/docs/release/3.19/manual/singlehtml/index.html#class-randomvariablestream>>.

# Security extension for live streaming in IP networks

*Martina Kesterová, Miroslav Michalko*

Department of Informatics, Faculty of Electrical Engineering and Informatics,  
Technical University of Košice, Slovak Republic

`martina.kesterova@cni.sk, miroslav.michalko@tuke.sk`

**Abstract** — The article deals with delivering video content over IP networks, with emphasis on content security and secure access. It focuses mainly on live streaming by means of HLS and RTMP. It also describes their core characteristics, such as the pros and cons of using these technologies, the content security concept and the handshake between client and server. It creates a system design for RTMP live streaming and HLS live streaming design. Authorization issues are solved by system user roles.

**Keywords** — streaming, live streaming, security

## I. INTRODUCTION

The development of technologies and increasing transfer speeds highlight the importance of multimedia content transmission technologies, which, however, have strict requirements for delay times. As data is generated in real time and must be received on time, the development looks towards technologies which are less and less tolerant of data loss and which are capable of delivering data content in the shortest time possible.

The critical issue of video streaming is that of secure IP flow access and content security. Access must be permitted to authorized personnel only. It is also necessary to confirm the real identity of a person and whether they are authorized to access these information. These issues are discussed in this article.

## II. GOALS

The article deals with video streaming system design and implementation in real time with emphasis on security. It describes the system in terms of component structure and streaming functionality. The system provides a user web-interface which enables streaming and watching live videos using the user's internal or external web-camera. Authentication and authorization security elements were added to secure access and data integrity. The described technology will be implemented on the Videosever portal of the Computer Networks Laboratory.

This task requires detailed streaming technology knowledge including the process of data cutting, assignment of identifiers, encapsulation method, handshake between client and server, encryption, authorization and authentication, data transmission and reverse processes.

## III. STREAMING TECHNOLOGY DESCRIPTION

There are two methods of video delivery over IP networks – progressive downloading and video streaming. Progressive downloading is a process which requires downloading the whole file from the server and storing it on the client's local disc before viewing. Streaming video content can be watched without downloading the whole file. The most common video streaming portals include YouTube and Vimeo.

The video streaming segment is immediately deleted after having been viewed. Video segments are continuously received and presented to the end user while being delivered by the provider.

According to the source [1], video streaming can be divided into two groups:

- live streaming,
- video-on-demand.

The term live streaming means content presentation to the end user in real time. A digital camera is an indispensable component of live streaming. In case of utilizing an analog camera, it is also necessary to use a digitizing device. Video-on-demand content is stored on a storage medium. In most cases, a server serves as a storage medium. When the client makes a request,

the stored content is displayed.

A streaming protocol is a technology responsible for data cutting, assignment of identifiers, encapsulation method, handshake between client and server, encryption, authorization and authentication, data transmission and reverse processes. Nowadays, [2] the most widely used streaming protocol is HTTP Live Streaming. RTP, RTSP, RTMP and RTMPE are other examples of streaming protocols.

#### A. RTMP Specifications

Real-Time Messaging Protocol was designed by the Adobe for high-performance audio, video and interactive content transmission between Adobe Flash Platform technologies. The RTMP protocol in parallel with the reliable TCP protocol guarantee time-stamped data delivering so the data is properly ordered at the end point device.

In the process of data transmission by means of RTMP, the server first cuts the streaming video into small pieces called *chunks*. Each chunk carries an identifier which is associated with the stream identifier the chunk was made of. During the transmission, the server has to wait until the whole chunk is delivered before sending a new chunk at the end point all chunks are grouped by their identifiers.

#### B. HLS Specifications

The HTTP Live Streaming Protocol was designed to work with the standard HTTP protocol without the need for special extensions, while it manages to ensure the same quality of service on both the client's and the server's side as any other protocol. Using this technology, the hardware encoder first encodes the input audio-video files – audio files are encoded into the H.264 format, video files are encoded into the AAC format. The output file is in the MPEG-2 Transport Stream format. The output file is subsequently cut by the streaming segmenter into small segments called chunks with the .ts extension. The standard segment length is 10 seconds. However, the HTTP protocol can operate with different segments lengths as well. The segments are stored in individual files and later on the web server. Streaming segmenter creates an index file called m3u8 containing a list of created segment files. The URL of the index file is published on the web server. The client's software reads the index and requests the files in corresponding order. The files are then displayed continuously.

### IV. SYSTEM DESIGN

#### A. System Component

The streaming video system consists of these components: multimedia content player, streaming sever and live streaming encoder.

Wowza Media System is a software solution providing video streaming. It supports live streaming and video-on-demand for all common players and devices. It enables configuration of a number of extensions, such as automatic video quality settings based on network traffic and client's hardware specifications, creating live streaming records along with the possibility of subsequent replay as video-on-demand, security and content protection, etc.

It is also necessary to choose a player that supports live streaming and video-on-demand via Wowza Media System, such as JW Player or Flowplayer.

A live streaming encoder is a streaming software solution that is capable of streaming audio-video input to the server in real time. The freeware version of Adobe Flash Media Live Encoder will be used as a live streaming encoder in this system.

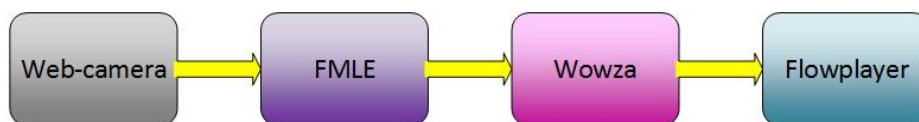


Figure 1 - Interaction of system components

#### B. Wowza Streaming Extensions

The Wowza Media System provides configuration for a number of extensions dealing with access security and content protection

StreamLock is an additional encryption function that provides a 256b certificate called Secure Sockets Layer, by which the customers who use the software services are verified. This method ensures high protection level by means of the RTMP protocol.

HTTPS or RTMPS deal with secure HLS, HDS or RTMP streaming using the SSL certificate. HTTPS or RTMPS do not secure data flow. They are used only in conjunction with SecureToken.

RTMPE is a service that is responsible for securing the RTMP streaming by creating an encrypted connection. It is often used in conjunction with SecureToken but it provides a lower protection level than RTMPS.

Digital Rights Management is a set of services that provides a system of key management and on-the-fly streaming content encryption. DRM is used in live streaming and also video-on-demand. In live streaming it provides the encryption method for each transmission separately with the possibility of changing the security keys.

SecureToken is a system that works on a request/response base. It is used for RTMP secure transmission. It is often used in conjunction with other services such as StreamLock, RTMPS or RTMPE. It provides a high protection level.

Wowza Media System also provides implementation of authentication for RTMP and RTSP protocol based on name and password.

The Hotlinking Protection service prevents the source code of the web page from being displayed and subsequently applied to another web page.

ServerSide API to Control Access is a secure service for controlling access to the streaming video content. It can be implemented for various protocols such as RTMP, Adobe HDS and HLS and Smooth Streaming. Together with secure transport services (such as StreamLock, SSL, RTMPS, RTMPE), it provides a high level of protection of streaming access.

Stream Name Alias is service that limits valid chunk names. It is also used for stream name mapping.

### C. System Design

The practical implementation of the problem is represented by a web site where the video player called Flowplayer is placed. Flowplayer has been chosen due to its compatibility with the Wowza Media System. The Wowza server provides a temporary repository for video files that are to be streamed to the end user via RTMP.

Authentication and authorization represent another important part of system. The methods available in Wowza are used for system security. The streaming platform has been extended by security modules that help protect the system, such as SecureToken, RTMP Authentication, Stream Name Alias and secure transmission via RTMPE.

Thanks to system scalability, the system can be expanded to mobile devices by means of HTTP Live Streaming which supports mobile device streaming.

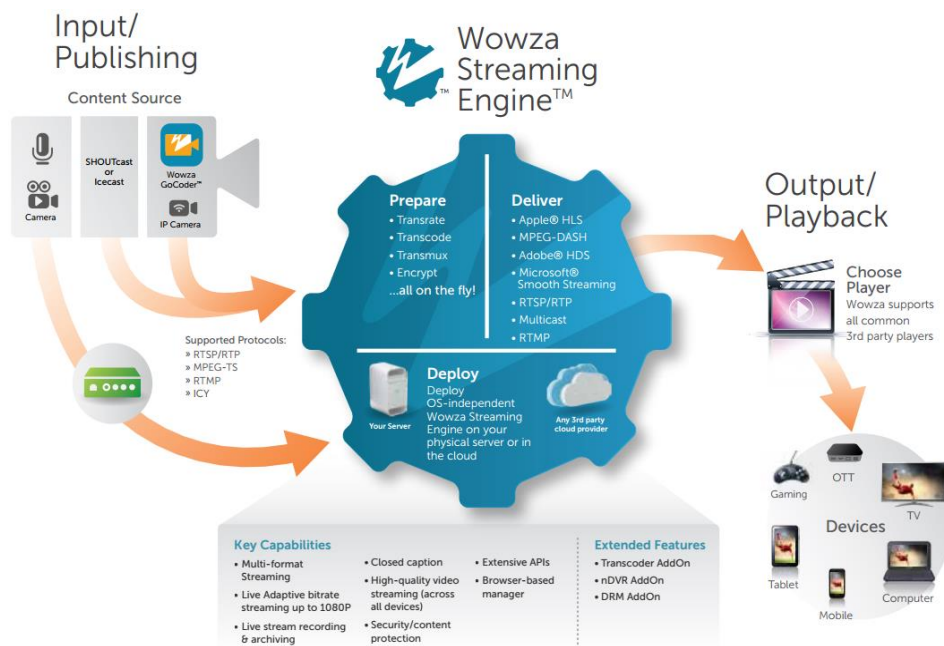


Figure 2 - Video streaming design [6]

*A. System Functionality*

On the server side, the Wowza Media System runs on a Linux platform. The system is not able to process client’s requests until appropriate ports are allowed.

The client side is independent from the operating system but on the other hand, an integrated or external web-camera is an indispensable component. Web-camera input is processed by Flash Media Live Encoder which connects to the Wowza server. If the connection is successful, the request for streaming can be created.

*B. Basic System Implementation*

A system implementation extending the secure streaming platform of Videoserver can be divided into several steps. First, it was necessary to make video streaming operational. This process includes choosing appropriate hardware/software components, installation of additional programs that are necessary for system operation, implementation of system functionalities on both client and server sides and their mutual connection. If the implementation is successful, there is need for system extension including security features that protect live streaming access. As the last step, the web-page has been created as a visible output of the system in which it is possible to watch live streams.

*C. Security Extension Implementation*

The following secure modules were implemented in the system:

- ModuleRTMPAuthenticate,
- ModuleSecureToken,
- OverridePublishLimitIP,
- system user authorization.

Using ModuleRTMPAuthenticate ensures requesting the name and password before the connection between client and server is established. The list of names and passwords is stored in a separate file.

ModuleOverridePublishLimitIP creates a list of specific IP addresses from which it is possible to watch the live streaming. An attempt to watch live streaming from an unlisted IP address is rejected.

SecureToken is a system providing a high protection level against spoofing attacks. It also protects the streaming video player. The both sides – client and server - share a mutual secret key and each connection is protected by another random key which is generated by server. This key is subsequently encrypted by the mutual secret key and it is sent to the client as a response to the client’s request. If the client knows the mutual key, they can decrypt it and send it back to the server. The server compares the sent value with the initial generated value. If it is a match, the connection is established. If not, the connection is aborted.

The creation of the web system solves authorization issues by system user roles. Each role has its own set of permissions. Each user who wishes to stream and watch live streams needs to be registered and signed up. The user will be given information for server connection which enables them to watch and stream videos in real time via their web camera.

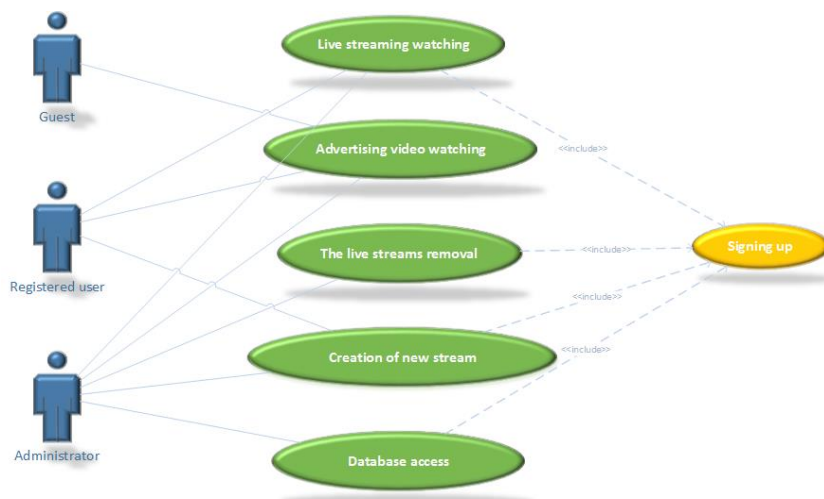


Fig. 3 - UseCase Diagram solving authorization issues

## VI. ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is financed from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## VII. CONCLUSION

The initial system has been implemented to analyze the current status of Videosever platform of Computer Networks Laboratory at the Technical University in Kosice. The task was to design a solution that will not require any change in existing streaming technology or cause a significant interference with the web interface. The article describes the design and implementation as well as an appropriate method of player authentication against the streaming server, so that the content had to be streaming from the local location and it was not possible to insert an embedded code to unauthorized web portals. In the implementation, two Wowza system supported players were considered. Due to its unique safety features, the Flowplayer software was selected. Another benefit of this article was analyzing the progressive technology of protection for the transfer of data between the streaming server and the client through HLS.

The task was to meet system requirements for video server security, availability and platform independence solutions. The resulting solution is deployed in test mode and its deployment into production solutions in the near future is planned.

## REFERENCES

- [1] JORDAN, Larry: The Basic of HTTP Live Streaming [online]. 2013. [cit. 2013-11-3]. <[http://www.larryjordan.biz/app\\_bin/wordpress/archives/2369](http://www.larryjordan.biz/app_bin/wordpress/archives/2369)>
- [2] ŽILINSKÁ UNIVERZITA. In: Scientific Letters of the University of Žilina. 2012, nr. 14.
- [3] IETF: 2013, HTTP Live Streaming draft-pantos-http-live-streaming-12.
- [4] APPLE Inc.: Streaming Media Guide [online]. Cupertino: Apple Inc.. [2013-08-08]. [cit. 2013-12-22]. <<https://developer.apple.com/>>
- [5] NIST: 2000, Advanced Encryption Standard
- [6] Wowza Media System [online]. Evergreen: Wowza Media Systems, 2014 LLC. [cit. 2014-5-21]. <<http://www.wowza.com>>

# Využitie platformy Arduino vo fotovoltaike

*Patrik ŠŤASTNÝ, Dušan MEDVEĎ*

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita  
v Košiciach, Slovenská republika

postastny@gmail.com, Dusan.Medved@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá využitím platformy Arduino pre aplikácie vo fotovoltaike, konkrétne problematikou vyhľadávania MPPT (maximálneho výkonového bodu) fotovoltaických panelov pomocou open-source platformy Arduino. S danou platformou bol vytvorený fyzický model trackera s riadením, ktorého algoritmus vyhľadávania MPPT bol porovnávaný s meraním bez použitia trackera.

**Kľúčové slová** — fotovoltaický panel, MPPT, Arduino

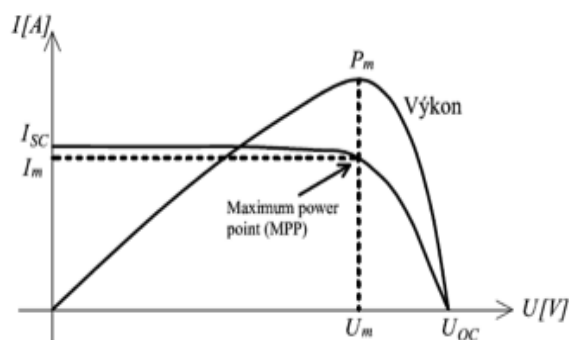
## I. ÚVOD

V dnešnej dobe je vývoj solárnych systémov zameraný predovšetkým na aspekty súvisiace s vývojom materiálov a procesov týkajúcich sa premeny slnečnej energie na elektrickú energiu. Pre lepšiu účinnosť celého systému sa využívajú polohovacie zariadenia (trackery), ktoré môžu zvýšiť energetický zisk až o 50 %.

Sledovanie bodu maximálneho výkonu (MPPT) je dôležitou časťou solárnych systémov, pretože pomáha zvyšovať efektivitu premeny slnečnej energie na elektrickú energiu. Všeobecne platí, že fotovoltaické články (FV) majú dva hlavné problémy a to účinnosť premeny elektrickej energie, ktorá je nízka (najmä pri slabej intenzite ožiarenia je to menej ako 17 %) a množstvo elektrickej energie generovanej solárnym panelom, ktoré sa neustále mení v závislosti od stavu počasia. Okrem toho, volt-ampérová charakteristika (VACH) solárneho článku sa v čase mení v závislosti od podmienok prostredia (osvetlenie, teplota). Na tejto charakteristike existuje jedinečný bod, alebo krivka, tzv. MPP, keď celý solárny systém (článok, prevodník, atď.), pracuje s maximálnou účinnosťou a vytvára maximálny výstupný výkon. Polohu bodu s maximálnym výkonom na VACH nepoznáme, ale jeho pozícia môže byť navrhnutá buď prostredníctvom výpočtových modelov, alebo nájdená pomocou vyhľadávacích algoritmov. Z tohto dôvodu sú MPPT techniky potrebné k udržaniu pracovného bodu na čo najvyššom výkone [1].

Sú vyvinuté mnohé metódy ako zabezpečiť sledovanie bodu maximálneho výkonu. Pri výbere vhodnej metódy sú rozhodujúce tieto kritéria:

- komplikovanosť riešenia,
- potrebné senzory,
- rýchlosť priebehu,
- cena,
- medza efektívnosti,
- zložitosť prevedenia,
- obľúbenosť.



Obr. 1 Výkonová charakteristika FV článku

## II. MODEL SOLÁRNEHO POLOHOVACIEHO ZARIADENIA

Model solárneho trackera je inteligentné polohovacie zariadenie, ktoré sa natáča za Slnkom prostredníctvom plne automatickej riadiacej jednotky, ktorá využíva riadiaci program pohybu Slnka, prostredníctvom časovača RTC DS1307 a dvoch modelárskych krokových motorov, vďaka ktorým je fotovoltaický panel vždy natočený smerom k najintenzívnejšiemu slnečnému žiareniu, tak, aby solárny panel vyrobil maximálne množstvo energie. Tracker je ďalej vybavený 2-riadkovým LCD displejom, ktorý užívateľa informuje o veľkosti výkonu a napätia dodávaného solárnym panelom. Ďalej zobrazuje aktuálny čas a udáva, v akom móde zariadenie momentálne pracuje. LED dióda signalizuje stav, keď solárny panel pracuje v MPP režime. Solárny tracker pracuje s platformou Arduino Uno R3. Pre svoju správnu činnosť musí byť tracker presne uložený v súradnicovom systéme (smerom na sever) podľa zadaného pomocného obrázka.



Obr. 2 Model vyrobeného solárneho trackera

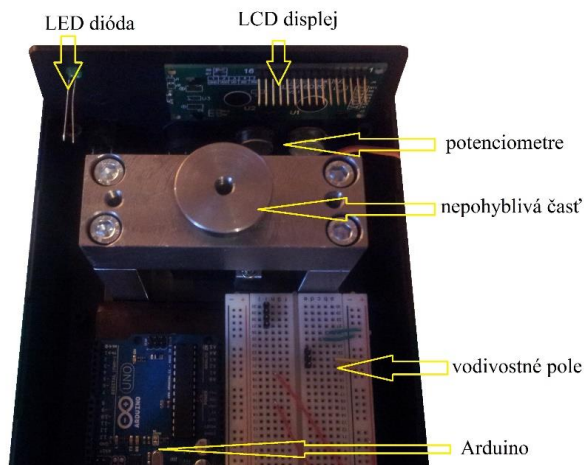
### **Technické parametre:**

- druh polohovania: dvojosí
- rozmery: 205 × 160 × 180 mm
- rozsah horizontálneho natáčania: 180°
- rozsah vertikálneho natáčania: 180°
- napájacie napätie: 9 V až 12 V
- výstupný výkon: max. 3 W
- riadiaca jednotka: platforma Arduino
- nastavovanie polohy: časovač, manuálne
- napájanie: 9 V batéria, USB port, 12 V externý zdroj



### Konštrukcia

Hlavná konštrukcia na polohovanie solárneho panela pozostáva z pohyblivej a nepohyblivej časti. Pohyblivá časť obsahuje krokový motor, ktorý otáča panel po horizontálnej osi a solárny článok. Táto časť je vyrobená z duralu. Je dostatočne pevná a zároveň ľahká, takže nezaťažuje vertikálny krokový motor pri svojej činnosti. Pevná, nepohyblivá časť, je zabudovaná v puzdre zariadenia a je vyrobená zo železa. Vďaka tomu je tracker ťažší a odoláva aj silnejším poveternostným podmienkam. Vo vnútri tejto časti je vstavaný krokový motor, ktorý natáča panel po vertikálnej osi. Podrobnejší pohľad je zobrazený na Obr. 3.



Obr. 3 Vnútny pohľad do solárneho trackera

### Ovládanie

Ovládanie solárneho trackera je sprostredkované niekoľkými ovládacími prvkami. Základ tvorí prepínač (MANUAL/AUTOMAT), ktorý dáva povel zariadeniu, aby pracoval v automatickom, alebo v manuálnom móde. Ovládanie v manuálnom prostredí zabezpečujú dva potenciometre, ktoré sú určené na polohovanie vertikálnej alebo horizontálnej osi. Automatický mód riadi vytvorený algoritmus. Tracker sa zapína, alebo vypína spínačom ON/OFF.



Obr. 4 Pohľad na predný panel ovládania

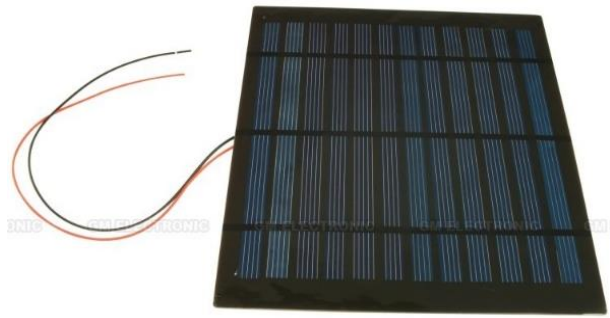
### Solárny panel

Na modeli trackera je použitý fotovoltaický článok značky OptoSupply, vyrobený z polykryštického kremíka (c-Si), ktorý je zapuzdrený v polymérplaste. Je odolný voči vode a UV-žiareniu. Takýto článok nedosahuje veľký výkon, preto sa používa len na rôzne experimenty a testovania, nabíjanie batérií, poprípade napájanie malých elektronických zariadení a LED svetiel. Jeho výstupné parametre sa líšia od klimatických podmienok, avšak produkuje energiu aj keď je oblačno, alebo je zamračené.

*Vlastnosti solárneho FV panela:*

- rozmery: 180 × 178 × 3 mm
- výstupný výkon FV panela (P<sub>MAX</sub>): max. 3000 mW
- výstupné napätie panelu (VOC): 6,0 V
- výstupný prúd FV panela (IMP): 500 mA

- odporúčaná pracovná teplota:  $-20^{\circ}\text{C}$  až  $+60^{\circ}\text{C}$
- počet kremíkových článkov: 48 článkov = 12 (články)  $\times$  4 (rad)
- tolerancia výkonu:  $\pm 3\%$



Obr. 5 Použitý solárny článok

**RTC DS1307**

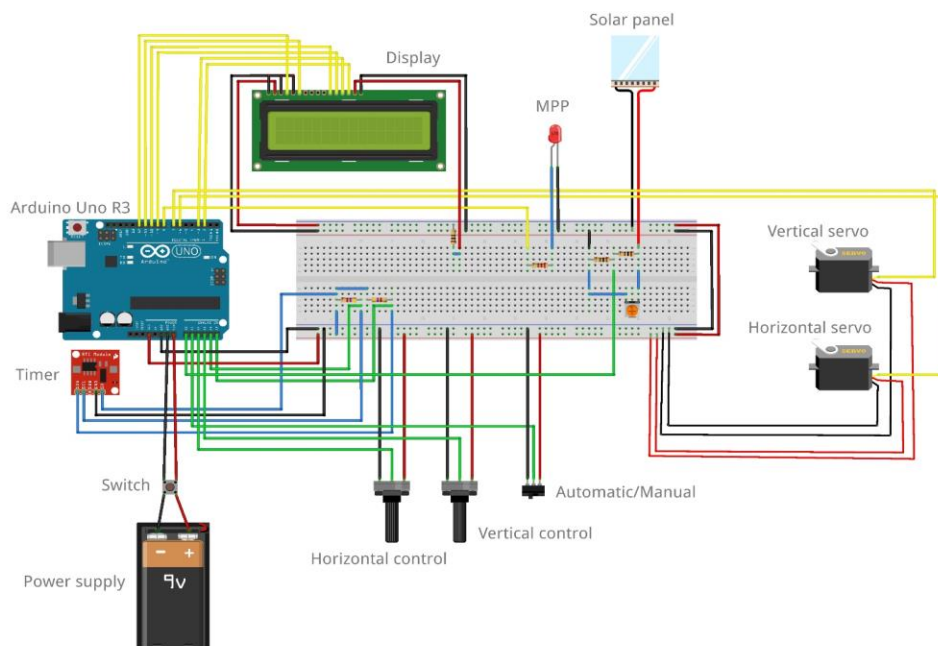
DS1307 serial real-time clock (RTC) je nízkonapäťový modul, plne binárno-kódovateľný časovač. Adresy a dáta sú prevádzané cez sériovú obojsmernú zbernicu I2C. DS1307 má vstavaný snímací obvod, ktorý sníma výpadok externého napájania a automaticky prepne na záložné napájanie, vďaka čomu nepretržite pracuje. Obsahuje lítiovú gombíkovú batériu, ktorá zabezpečuje chod časovača na minimálne 9 rokov (priemerne 17 rokov).

*Vlastnosti:*

- dva výstupy pre I2C rozhranie (SDA, SCL)
- hodiny, minúty, sekundy
- deň v mesiaci, dátum, rok
- kompenzuje priestupný rok
- presný kalendár až do roku 2100
- odporúčaná pracovná teplota  $-40^{\circ}\text{C}$  až  $+85^{\circ}\text{C}$
- 1 Hz výstupný pin



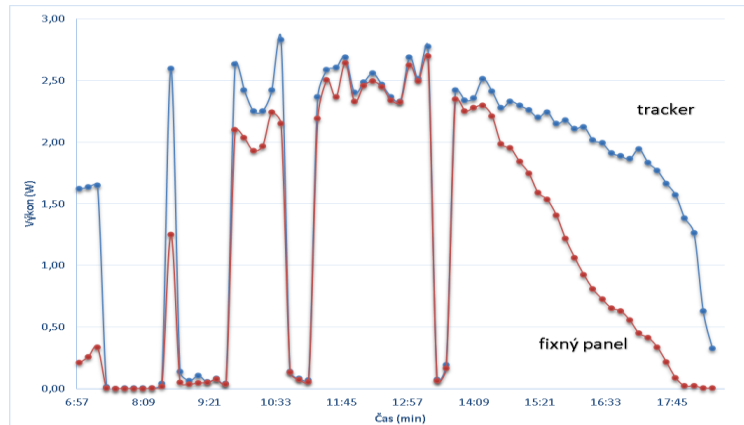
Obr. 6 Časovač RTC DS1307



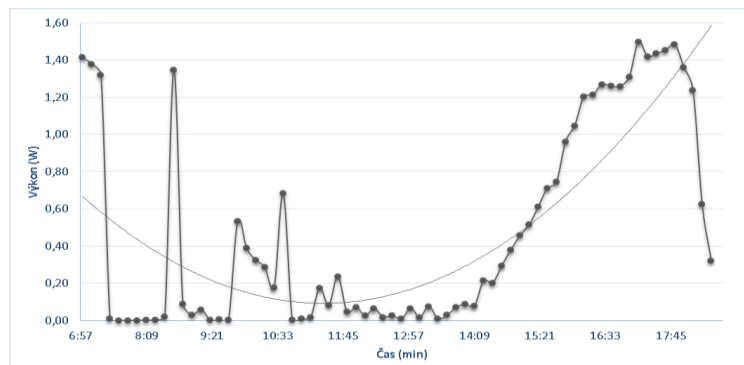
Obr. 7 Bloková schéma solárneho trackera

### III. VÝSLEDKY MERANIA

Aby bolo možné otestovať funkčnosť vytvoreného trackera a algoritmu, realizovalo sa celodenné meranie, ktorého výsledky sú uvedené na nasledujúcich obrázkoch.



Obr. 8 Závislosť výkonu na FV paneli  $P$  od času  $t$



Obr. 9 Závislosť rozdielu výkonu (fixný FV panel a tracker)  $P$  od času  $t$

Cieľom merania bolo zistiť elektrický výkon fotovoltaického panela počas dňa a energetickú účinnosť solárneho trackera. Údaje boli získavané z displeja solárneho trackera. Zaznamenával sa čas a napätie, z ktorého sa následne vypočítalo získané množstvo elektrickej energie. Merali sa dva stavy súčasne. V prvom stave bol solárny panel konštantne udržiavaný v polohe orientovaný na juh pri horizontálnom sklone  $45^\circ$ . V druhom prípade sa panel automaticky (podľa algoritmu) natáčal za Slnkom. Tieto stavy sa manuálne prepínali pomocou prepínača „MANUAL/AUTOMAT“ v pravidelných intervaloch 10 minút. Napätie  $U_{\text{FIX}}$  a výkon  $P_{\text{FIX}}$  v manuálnom stave sa líšia od veľkosti napätia  $U_{\text{T}}$  a výkonu  $P_{\text{T}}$  v automatickom móde.

Celodenné meranie sa vyznačovalo nestálosťou počasia. Intenzita žiarenia nebola konštantná. Na grafe (Obr. 8) vidieť razantné zmeny v dodávanom výkone, ktoré spôsobovala meniacia sa oblačnosť. Pri zamračenej oblohe bol energetický zisk takmer nulový. Priemerný nameraný výkon mal hodnotu pre statickú polohu  $P_{\text{FIX}} = 1,13 \text{ W}$  a pre tracker  $P_{\text{T}} = 1,61 \text{ W}$ . Výkonový pomer solárneho článku, navádzaného solárnym trackerom, je teda rovný:

$$P_R = \frac{1,61}{3} \cdot 100\% = 53,58\% \quad (1)$$

V porovnaní je to o 16,08 % viac ako pri výkonovom pomere statického článku:

$$P_R = \frac{1,13}{3} \cdot 100\% = 37,50\% \quad (2)$$

#### IV. ZÁVER

Cieľom tohto príspevku bolo prezentovať vytvorený algoritmus hľadania MPPT na zostavenom trackeri na báze Arduina. Model solárneho trackera slúži na demonštráciu fungovania polohovacieho zariadenia slnečných článkov, či kolektorov (v závislosti od použitia), ktoré potrebujú zlepšiť účinnosť premeny slnečnej energie na energiu elektrickú, prípadne tepelnú. Napriek dobrej účelnosti trackera by sa dal rozšíriť o mnohé moduly, ktoré by prispeli k lepšej vybavenosti polohovača. Požitím anemometra sa zaistí v prípade nepriaznivých povetnostných podmienok stav, kedy by sa tracker automaticky uložil do bezpečnej polohy, predišlo by sa prípadnému poškodeniu. Ďalšia z možností ako upraviť solárny tracker je pridať modul GPS. Ak je potrebné určiť presnú polohu Slnka na oblohe, potrebujeme vedieť údaje o aktuálnom čase a polohe zariadenia. Algoritmus by vďaka tomuto modulu automaticky aktualizoval potrebné dáta, bez potreby manuálne zasahovať do programu. Solárny tracker by tým pádom pracoval autonómne, prakticky kdekoľvek na svete. Pridaním Wi-Fi modulu sa zabezpečí lepšia komunikácia medzi zariadením a užívateľom. Užívateľ by vďaka tomu mohol na diaľku získavať informácie o výstupnom výkone, prípadne by vedel tracker sám ovládať. Ďalšia z možností je tracker vylepšiť o pamäť SD, ktorá by ukladala namerané výsledky. V neposlednom rade netreba zabúdať na solárny panel, ktorý sa dá vymeniť za výkonnejší, čím by sa dosiahla úplná energetická nezávislosť trackera od externého napájania. Použitie Arduino Uno R3 ale nedisponuje takou kapacitou analógových vstupov, aby sa pridalo ďalšie zariadenie. Výhodnejšie je použiť Arduino MEGA, ktoré je schopný pripojiť väčšie množstvo vstupných modulov (analógových, ako aj digitálnych).

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj pre projekt: Výskum charakteristík fotovoltaických komponentov pre efektívne projektovanie solárnych systémov, s ITMS kódom: 26220220080, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.



#### REFERENCIE

- [1] M. Berrera, A. Dolara, R. Faranda, and S. Leva: *Experimental test of seven widely-adopted MPPT algorithm*. IEEE Bucharest Power Tech Conference 2009. s. 1 - 8.
- [2] T. Esmar, - P. L. Chapman: *Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques*. IEEE transactions on energy conversion, vol. 22, no. 2, June 2007.
- [3] F. Liu, Y. Kang, Y. Zhang a S. Duan: *Comparison of P&O and Hill Climbing MPPT Methods for Grid-Connected PV Converter*. Industrial Electronics and Applications, 2008. ICIEA 2008. 3rd IEEE Conference, s. 804 - 807.
- [4] Šťastný P.: *Využitie platformy Arduino vo fotovoltaike. (Utilizing platform Arduino for photovoltaic applications)*. Bakalárska práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2014.

# WebRTC implementation for simple videoconference solution

*Emil HENČEK, Miroslav MICHALKO*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

emil.hencek@student.tuke.sk, miroslav.michalko@tuke.sk

**Abstract** — the aim of this work is the analysis as well as implementation of video conference system based on WebRTC technology. The application will enable users the ability to perform video calls without having to install application or any extensions to the current browser. Functionality of this video conferencing method is acquired only by means of a web browser. The introductory part of the paper focuses on the analysis of existing solutions, a description of how these solutions are functioning, description of their drawbacks and disadvantages which are to be eliminated with the use of this application. The paper also describes the technologies required to create a web application that would signal connection and synchronization. The core of the work is devoted to the implementation of mechanisms WebRTC. Furthermore, the implementation of the application follows. Final part of this thesis deals with testing and trial experiments of designed video conferencing solution.

**Keywords** — Web, video conferencing, WebRTC, HTML5

## I. INTRODUCTION

In today's world of digital communication, there are many applications to enable transmit video and sound in real time, without waiting. One of the most common applications providing this service is Skype. Skype and similar applications run in the client PC or mobile device. It is an individual application. The second group of these applications are web applications using a web browser to communicate.

Possibilities of video communications are constantly improving, because the evolution of browsers went in recent years rapidly ahead. Increasing the capacity of transmission media can enhance the quality of transmission. It is a hot topic because the web browser is a part of almost every computer and mobile device.

Web applications have many advantages. The main are that the user does not need your computer to install anything what often does not have permission. Application that runs in a web browser is independent of the operating system. If the browser works on the platform, a web application works too.

The goal of this [1] is to create a simple functional solution that enables the transmission video and audio between two or more browsers without the user having to install anything.

## II. ABOUT WEBRTC

**WebRTC** (Web Real Time Communication) is a Web standard that is supported by the latest versions of modern browsers and provides support for Web applications capable of audio and video calls independent on plugins. First implementation of WebRTC released Ericsson lab in May 2011.

WebRTC is a project with open source code developed by Google. The aim of this pioneering project is to provide web developers tools that facilitate the creation of applications for multimedia communications through a Web browser.

### A. Architecture of WebRTC

WebRTC offers web application developers the ability to write rich, real-time multimedia applications (think video chat) on the web, without requiring plugins, downloads or installs. Its purpose is to help build a strong RTC platform that works across multiple web browsers, across

multiple platforms. The overall architecture is shown on Fig.1.

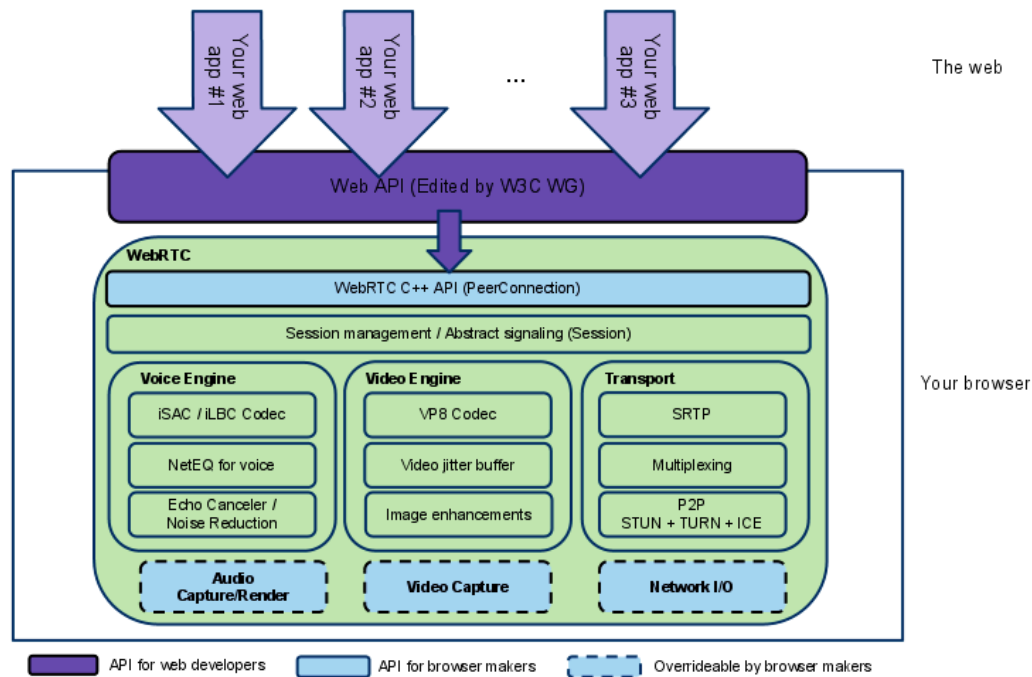


Fig. 1 – Architecture of WebRTC [2]

### B. WebRTC APIs

To acquire and communicate streaming data, WebRTC implements the following APIs:

**MediaStream:** get access to data streams, such as from the user's camera and microphone.

**RTCPeerConnection:** audio or video calling, with facilities for encryption and bandwidth management.

**RTCDataChannel:** peer-to-peer communication of generic data. MediaStream (alias getUserMedia)

## III. DRAFT OF VIDEO CONFERENCE SOLUTION

The solution was divided into three phases:

1. create a simple application that displays on the website stream from the camera,
2. develop applications able to make a video call within the same browser page,
3. resulting creation of functional application that is able to make a video call between any two (or more) computers on the planet.

Author puts final application the following requirements:

- simple
- quick
- no registration, filling in complex fields, setting
- no installation
- Web browser will do all the work

If the user A wants to connect with user B, so the user A enters into the browser and the web address of this service The application automatically generates the name of the room, which is basically a 32-digit code consisting of numbers and characters Is statistically impossible, so as to create two identical rooms Number of different rooms is so high that each cell of every person who lives on earth could have billions of their own rooms.

User A delivers the generated address to user B. He may do so by using any means of communication. It may be address sent by e-mail, via IRC channel, by telephone or other means it considers reliable.

User B enters into the browser address that it received. If both sides allow the browser to access the camera and microphone, connection takes place and both parties can communicate until either party terminates the current call.



#### IV. CREATION OF THE FINAL APPLICATION

The creation of the applications were used components described in the following subsection.

##### A. Components for creating applications

- **NetBeans** – IDE (Integrated development environment) with smart features,
- **XSocket.NET WebRTC JavaScript API** – javascript library, who simplify signalization,
- external signaling server using XSocet,
- **jQuery** – javascript library,
- **jQuery UI** – javascript framework, makes applications more interactive and visually more beautiful.

##### B. Structure of application

The application is located in several files:

- index.php
- main.js
- style.css
- functions.php

The main file **index.php** contains the basic HTML code and creates the basic structure of the page. It also takes inserted libraries that the application requires. The main set of JavaScript functions where the entire functionality of the application is in the file **main.js**. This file will also be dismantled further detail. The **style.css** file is stored Cascading Style Sheets, which give the appearance of the application. **functions.php** file contains functions that are needed to run the application, but their definition was necessary for clarity put into a separate file.

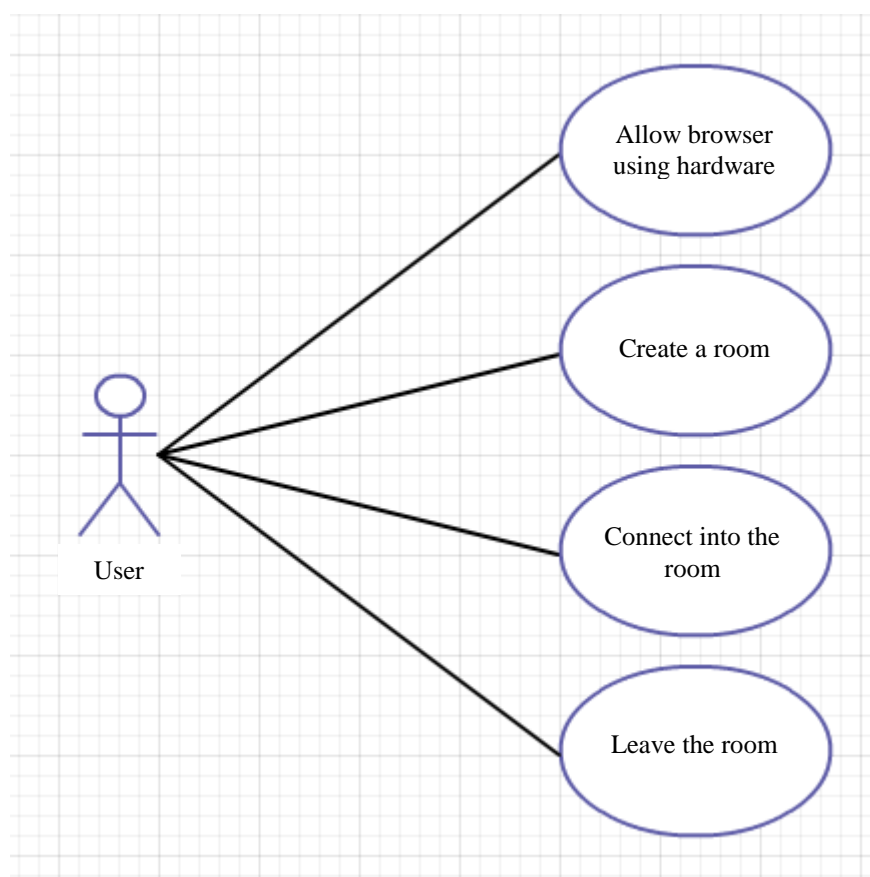


Fig. 2 - Use Case Diagram of implemented simple videoconferencing system

#### V. TESTING SOLUTIONS

The application has been tested at several devices in different browsers (Fig. 3). The following table shows the test results.

Table 1  
Results of testing application

Device	Browser	Functional
Asus K50IJ (notebook)	Chrome v. 34.0.1847.137	yes
	Opera v. 21.0.1432.67	yes
	Firefox v. 28.0	no
HP Pavilion DV7 (notebook)	Chrome v. 35.0.1916.114	yes
Prestigio Multipad 2 Ultra Duo 8.0 3G (tablet)	internal browser	no
	Chrome Beta v. 35.0.1916.114	yes
Samsung Galaxy S4 (smartphone)	Chrome Beta v. 35.0.1916.86	yes

The table shows that videoconference application works with all tested devices. Browsers, on which the application didn't run correctly, were the internal browser on the tablet with operating system Android and Firefox 28.0 on a laptop. The reason is that not all browsers have already implemented support for WebRTC. Firefox already has support for WebRTC. The application does not work because of JavaScript library that provides an interface between XSocket and WebRTC, which is not supported by Firefox yet. After complaints from users, we can expect a new version of the library will also support Firefox.

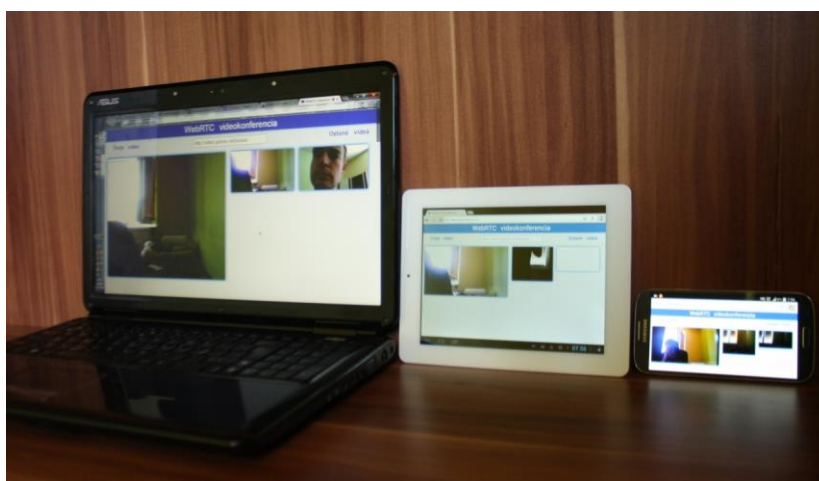


Fig. 3 – Final solution supporting various devices

## VI. CONCLUSION

The practical part of this thesis [1] is devoted to the implementation of videoconferencing solutions. Application is successfully implemented. People from all over the world can participate in video conferencing together with their friends in seconds. They do not need to install any application or add-on module. They just need to find the web browser with WebRTC. The application is available on the website: <http://video.gulraw.net>.

## ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is financed from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] Henček, E. Implementation of WebRTC for videoconference solution. Bachelor thesis. 2014. FEEL, Technical University of Košice, Slovakia
- [2] WebRTC architecture, Google Inc. 2013. Available online: <http://www.webrtc.org/reference/architecture>
- [3] jQuery API, The jQuery Foundation. 2014, Available online: <http://api.jquery.com/>
- [4] Web portal based solution designed for delivery and management educational multimedia content with auto conversion ability / Dávid Cymbalák, František Jakab, Miroslav Michalko - 2013. In: 19th Multimedia in Education : Conference Proceedings : June 13-14, 2013, Košice. - Kosice : Elfa, 2013 S. 29-32. - ISBN 978-80-8086-207-7



# EIGRP protocol analysis

*Egid HALUŠKA, Miroslav MICHALKO*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

egid.haluska@student.tuke.sk, miroslav.michalko@tuke.sk

**Abstract** — this article is focused on EIGRP routing protocol and its internal structure. Work presented by this paper is dedicated to analysis and possible findings of its weak spots and propose possible improvement. Afterwards EIGRP protocol was compared to RIP, RIPv2 and OSPF protocols. Main part of the work was focused on network experiments where behavior of dynamic routing protocols were monitored during simulated connection failures. Based on particular measurements, EIGRP protocol improvement in form of theoretical concept was suggested.

**Keywords** — EIGRP, RIP, RIPv2, OSP, convergence, dynamic routing

## I. INTRODUCTION

Computer networks are now part of our lives and became necessity. Networks have been evolving since early 70ies. In past, they were used mostly to simple purposes like browsing web pages or working with email. Data transmissions were not so heavily used and there was no such aim to have it 100 percent functional. Communication these days is changed and focus is on services like VoIP and video streaming. This implies also change of amount of transferred data. Connection outage can cause various complications. Remote data storages are popular now and contain various pieces of information to which we want to have constant access. These network connections are managed by bunch of routing protocols and it is important to focus on their functionality. They should be improved constantly since their main purpose is to recover after outage and to react in shortest possible time to various topology changes in network and, in case of change, find the best spare route as soon as possible. There exist various mechanisms for required time shortening.

Main aim of analysis was to introduce EIGRP network protocol. Describe how it works and what are its differences against other protocols and also to find its weak spots and emphasis strong ones. Therefore it was compared with numerous of other protocols implemented on devices in testing suite together with Cisco company devices.

## II. PROTOCOLS

### A. EIGRP protocol

Protocol uses DUAL algorithm, which is used for speed convergence routing optimization and also to avoid using of routing loops. Routing information is exchanged with neighbor nodes only once and then only updates are sent. If router is not able to find route via network, it'll send request to his neighbors, which then broadcast this request forward until appropriate route is found. This update is significant benefit comparing to other protocols, because it reduces route between routers and therefore also bandwidth is used more appropriate. Metric, which is used to find appropriate route is calculated via bandwidth, network load, delay and reliability. Incorporating this amount of variables into calculation ensures, that the best route is found. Comparing to other distance-vector algorithms, EIGRP has higher maximum value of skipping and thus it is compatible with big networks. Disadvantage of this protocol is that it is only compatible with Cisco components.

### B. RIP protocol

RIP is distance vector protocol type. It is one of the first protocols implemented to TCP/IP. Data part of RIP message is encapsulated into UDP segment. Each router, which is using this protocol, has limited knowledge about neighbor networks. This simple protocol uses mechanism, which counts number of skipping's between routers to find optimal route for packet. It uses 15 skipping's at most to avoid routing loop. This parameter limits number of networks supporting RIP protocol. However, this protocol is very popular because of its simplicity and configurability. It is best suited for small networks.

### C. OSPF protocol

OSPF is public routing protocol of link state type. This protocol routes packets of IP protocol collecting pieces of information about neighbor routers and building network map. OSPF broadcasts big amount of message types including contact messages (e.g. Hello message), requests for update and database description). In case that 2 cross-connected routers deal on common parameters, they become network neighbors. Adjacent routers send to each other LSU packets (Link-State Update), which contain advertisements LSA (Link-Stat Advertisement). Advertisements information describe router interface status or list of routers connected to given network. All routers stores LSA friends into local topology database and at the same time they send this information to other adjacent routers. Dijkstra algorithm is used to find shortest route and also to avoid network loops. In case that network topology is changed, router with changed topology broadcasts LSA and LSU packets to adjacent routers. These packets are then broadcasted via network and all routers will update their topology database and perform new calculation using Dijkstras algorithm. Advantage of OSPF protocol is its flexibility and determinism. In contrary to RIP, OSPF is too complicated which is its big disadvantage.

## III. CONVERGENCE

Traditional routing protocols of distance vector type (RIP, IGRP) uses periodical updates. Regarding unreliable nature of periodical updates, these protocols are vulnerable against network loops and endless calculations. RIP and IGRP use several mechanisms to avoid these problems including suppressing timer, which causes long time convergence. EIGRP doesn't use suppressing timers. Instead of that, routes without loops are calculated with different calculation system (dispersed calculation), which is executed coordinated manner. Between routers and their results is faster convergence as you can find in traditional distance vector protocol types [1].

## IV. METRIC

EIGRP protocol has a very useful attribute: on contrary to other protocols, which compare routes and choose optimal one based on one parameter, EIGRP uses combination of these 4 parameters:

- Bandwidth
- Delay
- Load
- Reliability

$$Metric = 256 * \left( \frac{K1 * 10^7}{BW_{min}} + \frac{K2 * BW}{256 - load} + K3 * \sum delays \right) * \left( \frac{K5}{reliability + K4} \right) \quad (1)$$

as far as  $K5=0$  the entire expression  $\frac{K5}{reliability + K4}$  is ignored

## V. MEASUREMENT OF CONVERGENCE RATE

### A. Test equipment and test conditions

The aim of this work was to find a weakness of EIGRP protocol and then propose improvements to what was achieved by comparison with other protocols. Based on theoretical assumptions, EIGRP protocol should converge the best, what was assessed using two test topologies. Tested topologies consist of two Cisco devices and two PCs that are involved according to the following schemes (Fig. 1).

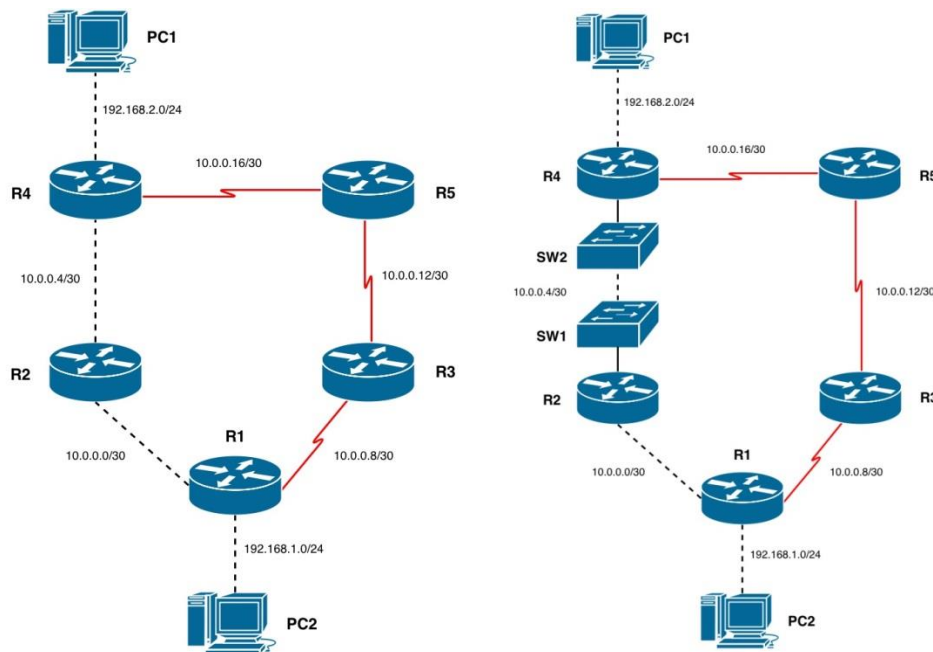


Fig. 1 Topology No. 1 and Topology No. 2

Individual devices were assigned different IP addresses for and so configured and connected devices gradually tested each protocol. Our convergence measurement was carried out so as to establish a communication between PC1 and PC2. Namely PC1 running under Linux operating system has been installed HPING3. That enabled us to regularly ping PC2 at various time intervals. Interval of 10 microseconds was chosen (10µs). So the ping from PC1 to PC2 was consigned 100 times per second. Pinging was done using the command: **sudo hping -1 192.168.1.2 -i u10000**, where 192.168.1.2 is the IP address assigned to PC2, which was ping-ed and -i u10000 is the time interval.

```

ubuntu@ubuntu: ~
len=46 ip=192.168.1.2 ttl=124 id=2095 icmp_seq=1198 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2096 icmp_seq=1199 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2097 icmp_seq=1200 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2098 icmp_seq=1201 rtt=10.5 ms
len=46 ip=192.168.1.2 ttl=124 id=2099 icmp_seq=1202 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2100 icmp_seq=1203 rtt=10.7 ms
len=46 ip=192.168.1.2 ttl=124 id=2101 icmp_seq=1204 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2102 icmp_seq=1205 rtt=10.5 ms
len=46 ip=192.168.1.2 ttl=124 id=2103 icmp_seq=1206 rtt=10.7 ms
len=46 ip=192.168.1.2 ttl=124 id=2104 icmp_seq=1207 rtt=10.5 ms
len=46 ip=192.168.1.2 ttl=124 id=2105 icmp_seq=1208 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2106 icmp_seq=1209 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2107 icmp_seq=1210 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2108 icmp_seq=1211 rtt=10.5 ms
len=46 ip=192.168.1.2 ttl=124 id=2109 icmp_seq=1212 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2110 icmp_seq=1213 rtt=10.5 ms
len=46 ip=192.168.1.2 ttl=124 id=2111 icmp_seq=1214 rtt=10.5 ms
len=46 ip=192.168.1.2 ttl=124 id=2112 icmp_seq=1215 rtt=10.6 ms
len=46 ip=192.168.1.2 ttl=124 id=2113 icmp_seq=1216 rtt=10.5 ms
^C
--- 192.168.1.2 hping statistic ---
1220 packets transmitted, 1217 packets received, 1% packet loss
round-trip min/avg/max = 10.4/16.3/1010.7 ms
ubuntu@ubuntu:~$ sudo hping3 -1 192.168.1.2 -i u10000

```

Fig. 2 Pinging using HPING3

Wireshark packet analyzer was installed on the PC2, through which was possible to capture the communication in the network. After running the program we set up the filter that captured ICMP packets from PC1. This filter ensured transparency of the packets in the program so the period of recovery could be quickly determined.

The program Wireshark captures every packet at the time when it was captured with microsecond precision, so this period could be determined. Therefore, during the period of connection failure was created time gap (which can be seen in Fig. 3 between the lines nr. 507

and nr. 540), when the program had not captured any packet so it was possible to determine the connection failure.

It was done by automatic export of tables with results from the Wireshark that was consequently opened in Excel, the program designed for work with tables. In this program, was the last time value sought that was captured before the connection failure and the first time value captured after the time gap created during the failure. After mutual deduction the time of the failure was found.

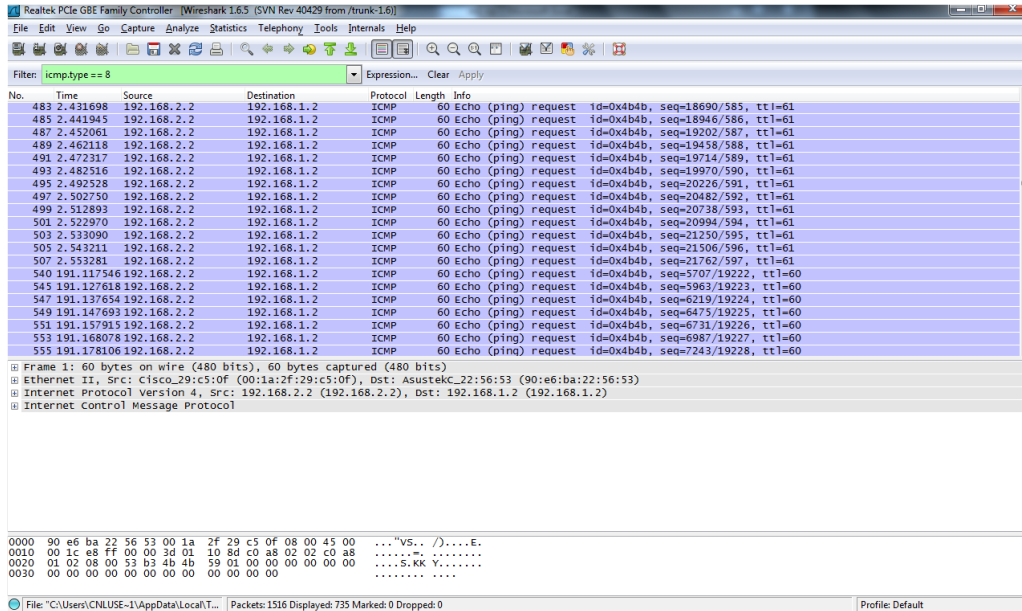


Fig. 3 Capturing the packets through Wireshark program

### Measurements were implemented on two topologies:

- Topology No. 1: After convergence of topology data were transmitted using the routers R4-R2-R1. Failure was simulated between routers R2-R4. During the failure was communication redirected to a route, which led through routers R4-R5-R3-R1.
- Topology No. 2: After convergence of topology data were transmitted along the route R4-SW1-SW2-R2-R1. Failure was simulated between SW1-SW2. During the failure was the communication redirected to a route, which led through routers R4-R5-R3-R1.

#### B. The measurements results

##### Topology No. 1

After testing it has been found that while failure of devices R 4-R 2 was reached the worst convergence RIPv2 in average 17.8473 sec. Protocols RIP and RIPv2 was unstable and between times were wide divergences.

Conversely, OSPF and EIGRP protocols had stable times of convergence. Best converging protocol was protocol EIGRP, for which the average time achieved 1.106507 sec. after adjusting the Hello timer to 1sec of the time value and Hold time to 3 sec. we have achieved an average improvement to 2.989678%, after which was the average time of convergence of EIGRP 1.073426 sec.

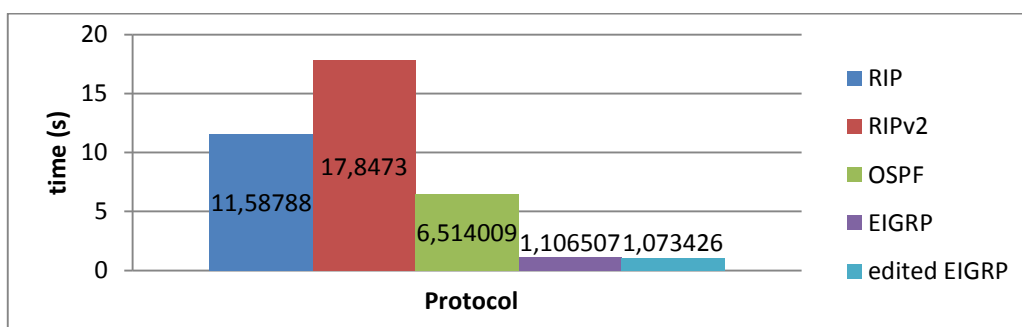


Fig. 4 The measurements results (Topology No.1)

## Topology No.2

By testing we found out that by the failure of devices SW2-SW1 was reached the worst convergence of protocol RIPv2 with in average 194.734 sec. Protocols RIP and RIPv2 behaved again unstable and between times were wide divergences. But protocols OSPF and EIGRP were also unstable, as in the first case, although the divergences were smaller.

Best converging protocol was again EIGRP protocol, which achieved in average 12.99257 sec.; after adjusting the Hello timer to 1 sec. and value of Hold time to 3 sec. was achieved an average improvement 80.509645%, after which was the average time of convergence of protocol EIGRP 2.532298 sec.

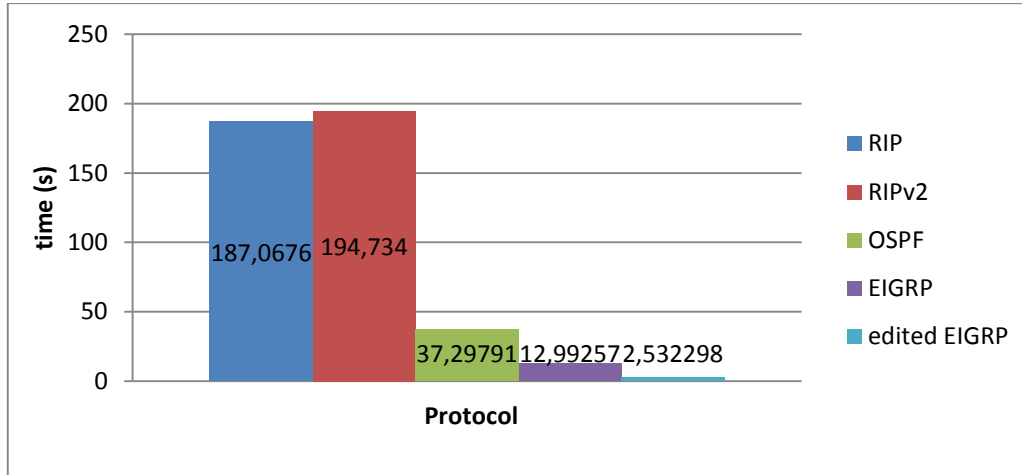


Fig. 5 The measurements results (Topology No.2)

Measurement has shown that in the topology no. 1 were all of the protocols converted faster. This rapid convergence was caused by failure simulation on the place where the network protocol was running while the network cable has been pulled out directly from the router. But in the topology no. 2 were also switches SW1 and SW2 placed before routers where the protocol was running.

The network behaves as there was a failure in real terms when using these switches. Measurement confirmed the theoretical assumptions that EIGRP protocols converges the best way from all of the protocols. The aim of this work was to suggest an improvements to protocol EGRP. Since this protocol reached in the convergence best results, there was an effort to improve the convergence, that was successful after adjusting the Hello timer value to 1sec and Hold time to 3 sec. The time of convergence on topology no. 1 improved for 2.989678% and on topology no. 2 for 80.509645%.

## VI. SUGGESTED IMPROVES OF EIGRP PROTOCOL

Measurement managed to reduce the time required for convergence of EIGRP protocol by reducing the Hello and Hold time times. However, this increased load and data services from the perspective of an increased amount of communication EIGRP packets. In our measurements were reduced to 1sec. values Hello and Hold time to 3 sec. which is three times of the Hello.

These values had to be reduced manually at each router separately. This solution is for the network management very difficult and time consuming because of the need to serve each router. Moreover, if there would be on the each router set values to 1 sec. at Hello and Hold time would be tripled Hello, would be occupied all lines by countless Hello and Hold time packets. This occupying in larger networks could lead to overloading of lines. It was therefore created a solution that is based on the fact that in each router would be installed the module that would automatically reduce these times. Reduce them according to the failures. After each failure on the each router would be the time reduced for 1 sec. at Hello and Hold time would be set at three times of the Hello packet. At the same time would the after each failure triggered preset time interval during which the module was able to further reduce the value of failure again Hello packet and Hold timer set to three times Hello. At the same time after each failure would run the preset time interval during which the module was able to further reduce the value of failure again Hello packet and Hold timer set to three times Hello. After each failure, the time interval would be restored and began to gradually decline. In case of a situation of the time interval reaches zero, the router should Hello interval increased by 1 sec and Hold timer would

be set at three times the Hello packet. Re-increasing intervals are important because if our modul would decrease the values after each failure of Hello and Hold times then switches would occupy the lines Hello and Hold time by packets and the line would be overload Therefore, if during the time interval there was no failure, only times of Hello and Hold timer have been increased.

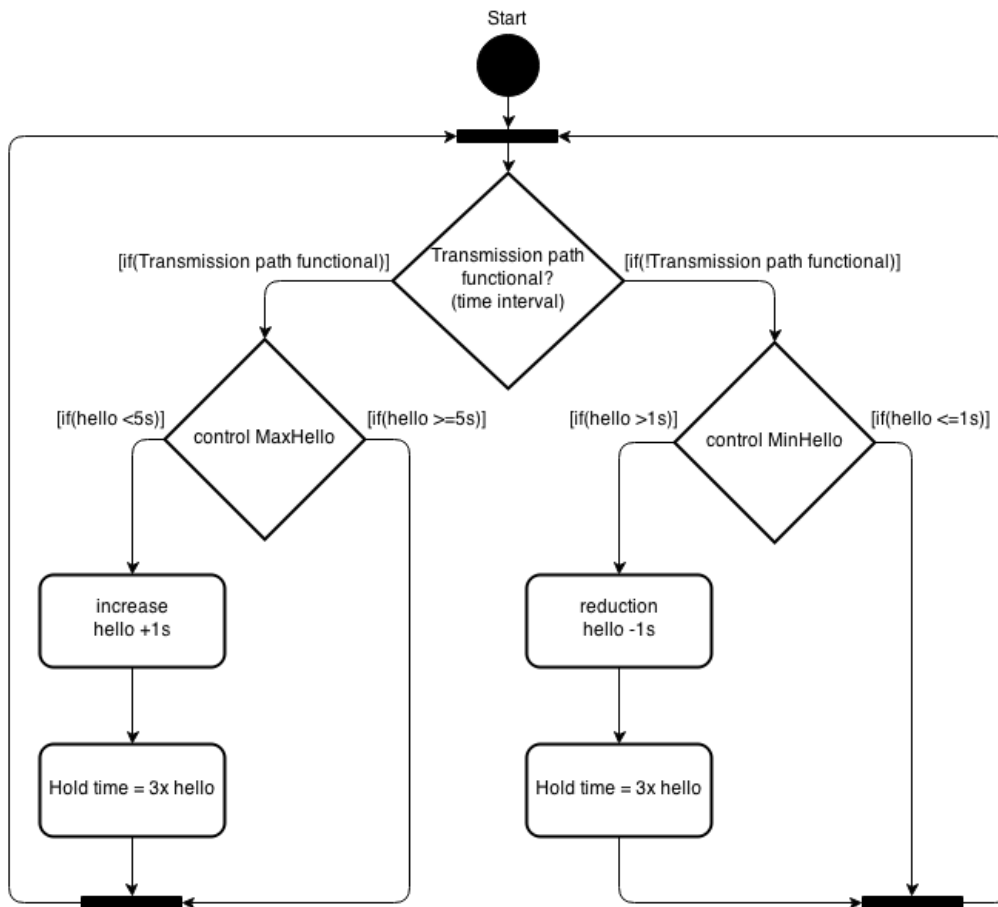


Fig. 6 The algorithm proposed solution

## VII. CONCLUSION

Papers presents advantages and weak points of EIGRP protocol in compare to other dynamic routing protocols based on real world scenarios. Topology and measurement were realized in laboratory conditions as presented above within Computer Network Laboratory at Technical University in Kosice. Paper offers algorithm which could improve behavior of EIGRP on unstable link between two routers. This work is just a basis for further development and measurements.

## ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is financed from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] CCNA Exploration: Směrování, koncepce a protokoly. Plzeň: VOŠ a SPŠE Plzeň, 2012.
- [2] ANTHONY, Bruno – JACQUELINE, Kim: CCDA® Exam Certification Guide (CCDA Self-Study, 640-861), 2nd Edition. Indianapolis: Cisco Press, 2003. 303-304s. ISBN: 1-58720-076-7
- [3] LAMMLE, Told: CCNA Výukový průvodce přípravou na zkoušku 640-802. Brno: Computer Press, a.s., 2010. 422 - 430s. ISBN 978-80-251-2359-1
- [4] EIGRP deep protocol analysis. Egid Haluška. 2004. Bachelor work. FEEL, Technical University of Košice.

# Extension of smartphone functionality using external sensors

Igor ČOLLÁK, Miroslav MICHALKO

<sup>1</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

igor.collak@student.tuke.sk, miroslav.michalko@tuke.sk

**Abstract** — this paper is deals with integration of various external sensors into Android powered devices using Arduino ADK as external extension to mobile phone. Goal was to demonstrate potential of such devices to deliver new range of services to mobile end users. Solution presented within this paper is polygraph device with human body sensors and software application for mobile devices which shows trustworthiness of answers. Paper is mainly focused on technical aspects of such solution, not on calibration of body parameters followed by presented software application. Result is part of work under guidance of Computer Networks Laboratory at Technical University in Kosice.

**Keywords** — smartphone, sensors, polygraph, Arduino

## I. INTRODUCTION

According to predictions proclaimed by top world institutes or overall global market trends, in the future the smartphone will dominate as a handheld communication device. General trends are clear and directions are defined – go mobile. Many businesses are now formed and established businesses are extending into new prospective fields such as smartphone market and related applications ecosystems. In May 2013, one of the leading companies operating one of the most extensive ecosystem for mobile applications announced that 50 billionth application has been downloaded [1] [2].

## II. THEORETICAL SMARTPHONE EXTENSION

Smartphones indeed represent state-of-the-art technology which is available to ordinary people, newest models of smartphones have implemented a number of different technologies and functions, e.g. camera with autofocus, display, loudspeakers, microphone, GSM module, GPS module, proximity sensor, accelerometer sensor, gyroscopic sensor, Wi-Fi module, Bluetooth module, and Memory module.

All these integrated technologies make smartphones more powerful and more usable like old way personal computer; in addition, they often have much better hardware performance compared to standard computers. For good demonstration of usability of compressed technology into small handheld devices we can mention some applications which effectively benefit from this hardware source.

There are useful applications which are using GPS module and thus provide navigation services, other applications use high resolution camera and OCR technology for detecting and labelling real locations from camera preview or directly translating words focused with camera. There is also, for example, a possibility of recording human voice and use it as security identification; this approach has been used in several applications, too. So possibilities of using smartphones are almost infinite, but despite these facts in some situations specific action cannot be achieved.



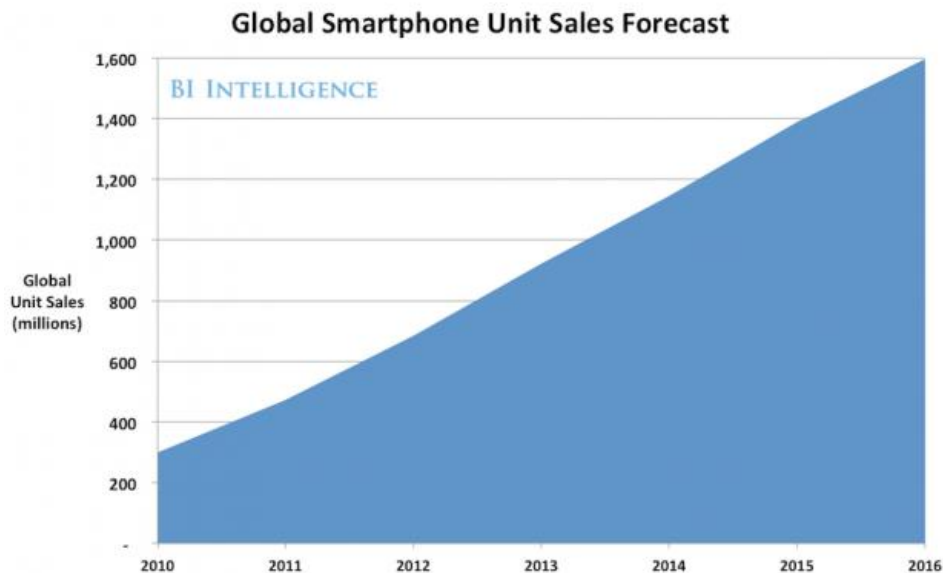


Fig. 1 Smartphone sales forecast [1]

One such situation is described in this paper, the goal is to construct the sensor set attached to smartphones with Android operating system. This sensor will actively measure given parameters from the ambient environment, thus extending functionality of the Android device.

In order to solve this problem extending smartphone about external sensors we have suitably chosen the polygraph instrument, whose nature and function very well demonstrate just that.

It was decided to make following measurement of physiological functions: heart rate, frequency of breathing, and galvanic skin resistance. These three physiological functions are necessary in any professional polygraph device and usually there are two additional measurements: blood pressure and frequency of breathing measured at belly.

Firstly, this instrument which has undergone tremendous development and developed into its present form will be described.

### III. ANALYSIS OF POLYGRAPH DEVICE AND ITS USABILITY FOR EXTENSION SMARTPHONE ABOUT ADDITIONAL FUNCTIONS

Polygraph device becomes a frequently used tool when seeking the truth, many people know this device by the popular name “lie detector”. This title could make a wrong impression that the device is able to detect whether somebody is lying, actually this is not true because polygraph can only measure physiological processes, and based on the evaluation of those measurements, the examiner can deduce suspicion that someone is lying. Even though the polygraph cannot detect lie directly, it is a powerful tool for solving criminal cases, principle of the device lies in the fact that no one can lie about something serious without at least minimal alteration of physiological processes of his body. The fact is that with the increasing knowledge of physiological processes and biochemical functions of the human body and also with rapidly perfecting technique, supported by mathematical analysis and comparisons - polygraph increases reliability of results. Recent research says that the accuracy of the newest computer polygraphs is nearing 100%.

### IV. TECHNICAL ASPECTS OF THE SOLUTION

As is apparent from the description in section three polygraph plays an important role in determining the truth in different areas, in this section we will focus on some technical aspects of construction such a device. To achieve our given goal we will use the following technologies and hardware, first we use the already-mentioned Android smartphone, the second major side of our device under development represents evaluation board Arduino ADK Mega2560, for connecting these two platforms we will use USB interface and ADB (Android Debug Bridge)



protocol which allows wide compatibility and supports also older models of android devices.

Most time consuming part of construction that polygraph device, is writing Android application for this has been chosen IDE (integrated development environment) Eclipse program which is a very powerful tool for developing Android applications with a wide range of functionality like graphic user interface editor or advanced debugging options and many others. For developing software for Arduino development board we use Arduino software development studio.

Regarding principle of sensors as transformations from physiological domain into electric signals we use three customary approaches, each for a single sensor. Measurement of heart rate is realized with usage of infrared diode and transistor, principle lies in the fact that infrared light is partially reflected from human blood so when infrared diode and infrared transistor mounted side by side and facing the same are attached to a finger, the light emitted from diode is reflected from finger and detected by infrared transistor so when transistor is 'open' thanks to reflected light and current flows through it, small alterations could be observed in electric signal in voltage level so after magnification and filtration of that small peaks the improved output signal constitutes the heart rate.

GSR (Galvanic Skin Resistance) circuit uses an operational amplifier MAX4251 in connection with negative feedback, the voltage applied to the resistive divider of voltage is discharged through a  $1M\Omega$  resistor connected to the ground to which are connected in parallel also terminals, what are during operations firmly attached on the fingers of the person. In the case that during the test there is a change of skin resistance due to sweating, a reduction in resistance in the total resistance of the underside of the divider decreases, which results in a decrease in the voltage supplied to the positive input of the amplifier and hence a significant decrease in its output. To invert the change from a decrease to an increase one PNP transistor has been used, so the resulting chart shows the rising value of the signal at increased sweating.

Respiratory rate sensing circuit is very simple, containing one active element NPN transistor. To base of the transistor is connected grounding  $100K\Omega$  resistor, in this joint is also connected a runner pin of  $10K\Omega$  potentiometer, the remaining two pins of potentiometer are in typical connection first one is grounded and other pin is fed with positive voltage through a  $100\Omega$  resistor. Potentiometer itself is sewn into the belt with his runner adapted the way so that when the tested person inhales and exhales, runner of potentiometer is moving.

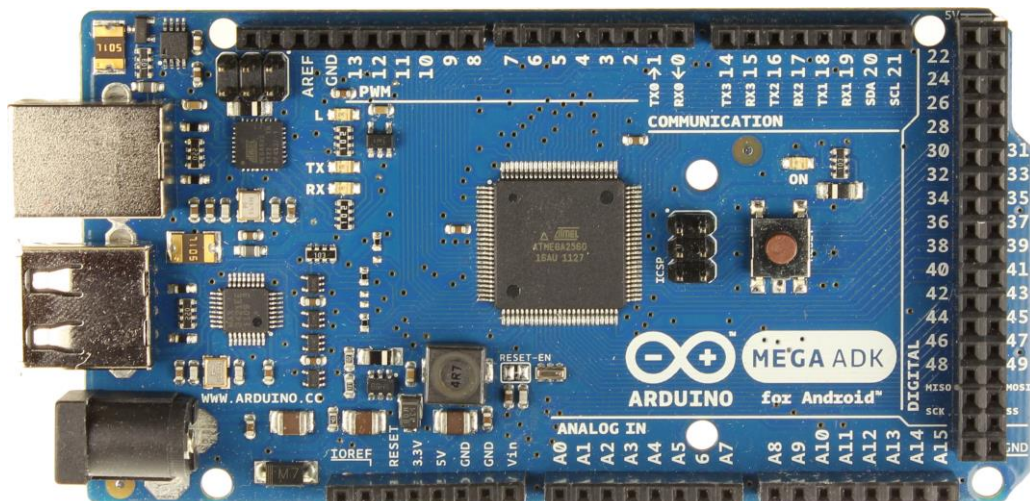


Fig. 2 Arduino ADK Mega2560 0

## V. DESCRIPTION OF THE OPERATION AND SUMMARY OF THE RESULTS

Polygraph device described in this work simulates operation of professional polygraphs, the main component is Android devices which serves as display unit. Android application receives data which are processed across three separated Java threads, purpose of the first thread is to establish connection between Android device and evaluating board Arduino ADK Mega as shown on Fig. 2.

The principle of operation is known network communications model client – server, on the computer (Android device) that communicates with the appliance running server that controls communication so that it accepts commands from the client, which is also an instance on a computer (Android device) and sends these commands to the device serially, thus phone with Android OS where is running daemon instance, is able to receive and send commands.

Second major Java thread is called Draw\_Engine as name suggests in this thread are data processed, evaluated also there are placed key methods for data comparison and possible lie evaluation. The main point of this thread is to lighten processing requirements for main graphical thread which then directly represents given data.

Third main Java thread acts as monitor in meaning of direct representing of input data. For random signal drawing was used Open GL ES library which perfectly suits for this situation, however this powerful library was used just for signal animation and remaining graphic are elements available from Android application development. Final main page of Android application contains intuitive symbols, in top of screen are illustrated lie and true symbols. On the left side are located symbols of the measured waveform, and at bottom of the screen are four buttons for setting and execution. Final application screen is illustrated on Fig. 3.

Algorithm to obtain the heart rate monitor works as follows: during the active measurement is created and started an instance of timer serving as time reference. In the body of the timer is periodically (10-ms) incremented monitored variable of time. While timer is running and during testing, the input values of heart rate electric signal is recorded and stored in a field, in every iteration is compared the previous value against current value and if the current value is greater than previous value for about by condition specified number of times the rising edge of the single heart beat is found and block of code in body of condition is executed: store the current value of the time of rising edge to an auxiliary variable serving as a comparison. As the program runs, continue to seek further rising curve of another heart beat and also upon its finding is saved value of the variable time, after fulfilling the conditions of finding a specified number of rising edges of beats block of code within condition is executed which actually compares the growth of times and calculates average value what results as approximate frequency heart rate per minute. Other algorithms works similarly, such the detector of changes the resistance of the skin as well as respiratory frequency meter, results of these measurements are collectively evaluated to determine whether a person is lying or not. Polygraph device which was successfully constructed can record variations in the measured of physiological functions but for real deception detection with sufficient reliability is required further development. PCB is shown on Fig. 3.

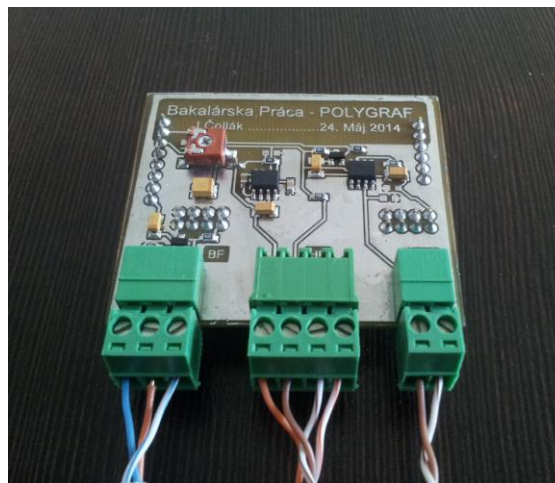


Fig. 3 PCB of Polygraph implementation

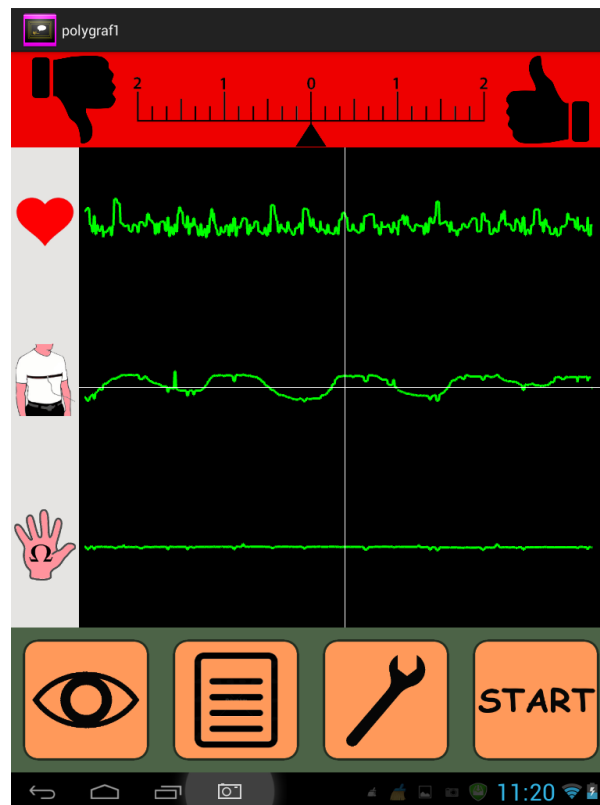


Fig. 4 Android application Polygraph UI

## VI. CONCLUSION

Papers presents cooperation on Android based mobile device and Arduino board to extend functionality for end user. Design and integration shows potential of Arduino boards to deliver multiple data sources to mobile phone which could make calculations and graphical analysis. Implemented polygraph tool offers interesting functionality but due to complexity of this field it is not practically implementable on market. Anyway this implementation shows potential for future development of solutions for medicine or smart homes.

## ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is financed from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] PASCAL-EMANUEAL GOBRY - ANALYST: Smartphone Sales Will Dwarf PC Sales This Year And Reach A Staggering 1.5 Billion Per Year By 2016 [online]. 2. February 29, 2013. [cit. 2014-05-03]. Available on internet: < <http://www.businessinsider.com/smartphone-salesforecast-2012-2>>
- [2] INFORMATIVE FACTS, CASE STUDIES, CHARTS & VIDEOS – Case Studies [online]. [cit. 2014-03-02]. Available on internet: <http://www.polygraphis.com/casestudies.htm#platinum>>
- [3] Arduino – Arduino MEGA ADG 2560 [online]. [cit. 2013-11-28]. Available on internet: <http://arduino.cc/en/adg-2560>
- [4] Polygraph accessory device for Android smartphone. Igor Collak. 2014. Bachelor work. FEEL, Technical University of Kosice.

# Teplotná analýza elektrických kontaktov vysokoprúdových zariadení

Tomáš POLÁK, Dušan MEDVEĎ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

tomas.polak@minet.sk, Dusan.Medved@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá analýzou kontaktov pre vysokoprúdové zariadenia. Problematika elektrických kontaktov je oblasť, v ktorej sa stretávajú úlohy riešenia materiálových, mechanických, elektrických a v neposlednom rade teplotných vzťahov popisujúcich procesy prechodu elektrického prúdu kontaktami. Poznanie presného rozloženia teplotného poľa elektrických kontaktov a v ich okolí, umožňuje výrobcovi správne zvoliť materiál, nadimenzovať tvar, prípadne upraviť vodivú cestu tak, aby sa elektrický kontakt neprehrial a plnil funkčnú úlohu opätovného spínacieho prvku v danom elektrickom obvode.

**Kľúčové slová** — teplotné pole, elektrický kontakt, ANSYS, vysokoprúdové zariadenia

## I. ÚVOD

Každý elektrický spínací prístroj musí mať prúdovú dráhu usporiadanú tak, aby sa dala v jednom mieste počas prevádzky prerušovať. Miesto, v ktorom dochádza k prerušovaniu, ale aj k prepojeniu dráhy, vykazuje oproti ostatným úsekom niektoré rozdielne vlastnosti. Vodivé spojenie dvoch oddeľovaných častí prúdovej dráhy sa realizuje pomocou kontaktov.

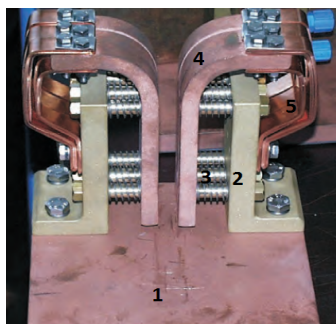
Prerušovanie elektrickej prúdovej dráhy sa môže uskutočňovať za rôznych prevádzkových podmienok, napr. s prúdom, bez prúdu, pri rôznych napätiach, s veľkou hustotou spínania a pod. Táto skutočnosť zaraďuje kontakty medzi dôležité časti prístrojov. V súčasnej dobe sa kladú na kontakty vysoké požiadavky a preto sa im venuje značná pozornosť. Všeobecne môžeme požiadavky kladené na kontakty charakterizovať takto [1]:

- a) malý stykový odpor,
- b) veľká odolnosť proti mechanickému opotrebovaniu (hlavne proti treniu kontaktov),
- c) veľká odolnosť voči zváraniam,
- d) veľká odolnosť voči opaľovaniu elektrickým oblúkom,
- e) optimálny vplyv na vývoj deionizačných pochodov po uhasnutí, vypínacieho oblúka (dobrá spínacia schopnosť).

## II. RIEŠENIE TEPLOTNÉHO POĽA ELEKTRICKÉHO KONTAKTU

### *500 A elektrický kontakt od spoločnosti Druseidt*

Ako príklad výpočtu rozloženia teplotného poľa bol vybraný 500 A elektrický kontakt od spoločnosti Druseidt, ktorý je variabilný a je možné podľa výrobcu upravovať niektoré jeho rozmerové alebo materiálové parametre.



Obr. 1 Hlavné časti predmetného kontaktu (1 – základná doska; 2 – pomocne držiaky; 3 – prítlačné pružiny; 4 – sedlové kontakty; 5 – nôž, ktorý zapadá do sedla)

#### ***Použitý materiál jednotlivých častí:***

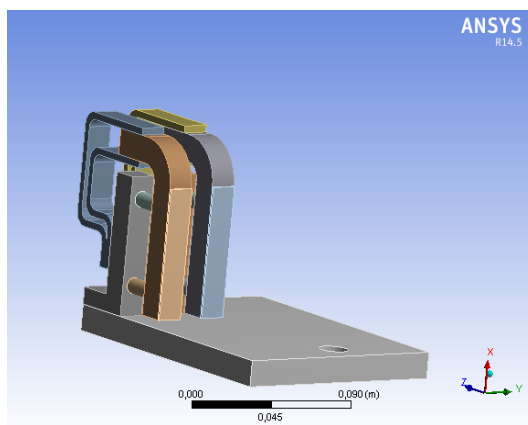
- základná doska je z medi,
- pomocné držiaky sú ocelové,
- prítlačné pružiny sú ocelové,
- sedlové kontakty sú z leštenej medi,
- nôž je medený.

#### ***Rozmerové parametre nečistoty pri modelovaní***

Pri vytváraní modelu nečistoty bola použitá hrúbka nečistoty 1 mm, hoci v skutočnosti môže dosahovať hrúbku aj  $10^{-8}$  m kvôli zjednodušeniu vysieťovania. Tomu boli prispôsobené aj elektricko-teplotné parametre tejto vrstvy.

#### ***Geometrické parametre spoja:***

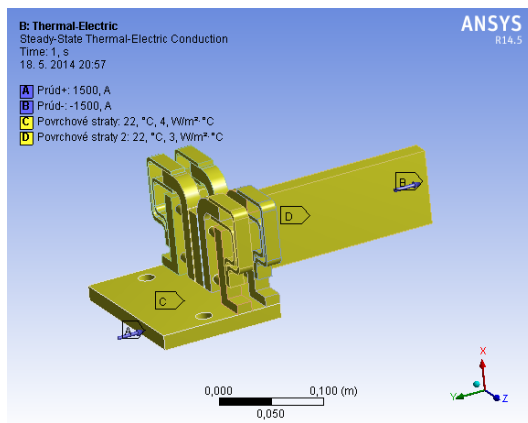
- výška jedného spoja: 75 mm,
- šírka jedného spoja: 15 mm,
- hrúbka nečistoty: 1 mm,
- plocha 1 spoja:  $1125 \text{ mm}^2$ ,
- celkový povrch nečistoty:  $4860 \text{ mm}^2$ .



Obr. 2 Zobrazenie plôch, na ktorých boli uvažované nečistoty

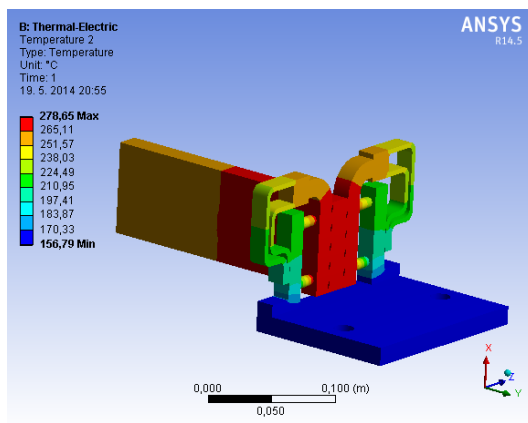
#### ***Tepelné povrchové straty***

Pri modelovaní musíme brať do úvahy aj miesto, kde sa daný kontakt používa. Pre výpočet bolo potrebné stanoviť okrajové podmienky, teda povrchy stien, z ktorých sa uvažuje s prestupom tepla do okolia. Pre horné steny kontaktu bol koeficient prestupu tepla  $\alpha$  zvolený  $4 \text{ W}\cdot\text{m}^{-2}\cdot\text{K}^{-1}$  a pre bočné a spodné steny  $3 \text{ W}\cdot\text{m}^{-2}\cdot\text{K}^{-1}$ .

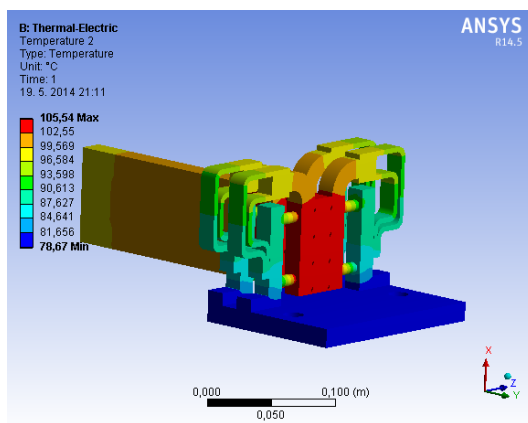


Obr. 3 Zobrazenie okrajových podmienok na jednotlivých miestach kontaktu

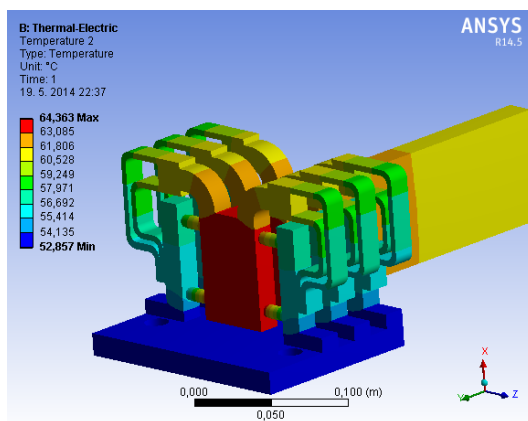
Pri analýze teplotného poľa bolo potrebné zistiť, ktoré miesta sú najviac tepelne namáhané a teda bolo potrebné stanoviť, minimálnu stykovú plochu a teda počet sediel. Jednotlivé alternatívy pri použití 1-, 2- a 3-sedlových kontaktov sú uvedené na nasledujúcich obrázkoch.



Obr. 4 Teplotné pole 1 páru sedlových kontaktov



Obr. 5 Teplotné pole 2 párov sedlových kontaktov



Obr. 6 Teplotné pole 3 párov sedlových kontaktov

Pri použití troch párov sedlových kontaktov je plocha styku  $9000 \text{ mm}^2$ . Najvyššia teplota je v mieste styku,  $64,363 \text{ }^\circ\text{C}$ .

### III. ZÁVER

Tento príspevok mal za úlohu predstaviť analýzu teplotného poľa vysokoprúdových elektrických kontaktov, ktorá bola realizovaná v prostredí ANSYS. Z porovnaných výsledkov je zrejmé, že využitím väčšieho počtu sedlových kontaktov a teda zväčšením dotykovej plochy sa zníži prechodový odpor kontaktnej vrstvy a tým sa znížia Joulove straty a teda aj teplota samotného kontaktu. Celkový prúd sa tak preteká väčšou plochou a menej zaťažuje niektoré úzke miesta, ktoré by sa prehrievali. Ďalšou alternatívou k zväčšovaniu stykovej plochy je zmena materiálu, úprava materiálu, ako aj zvýšenie prílačnej sily kontaktov, poprípade doplnenie chladenia (aktívneho, pasívneho).

### POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied grantom VEGA č. 1/0388/13.

### LITERATÚRA

- [1] HAVELKA, Otto a kolektív: *Elektrické prístroje*. Praha SNTL/Alfa 1985.436s.
- [2] Druseidt elektrotechnik: *Product information 02/2011*. Dostupné na internete: < <http://www.druseidt.de/pdf/0211eng.pdf> >.
- [3] *Contact-Systems and Accessories for Anodising and Electroplating Equipment*, 2000, Dostupné na internete: < <http://www.druseidt.de/pdf/kat3eng.pdf> >.
- [4] POEÁK, Tomáš: *Teplotná analýza kontaktov pre vysokoprúdové zariadenia*. (Thermal analysis of contacts for high-current devices). Bakalárska práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2014.



# Application for monitoring of the CNL Videoserver platform

*Filip BALINT, Miroslav MICHALKO*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

balint.filip@gmail.com, miroslav.michalko@tuke.sk

**Abstract** — This paper deals with monitoring of selected parameters of Videoserver streaming platform. This platform is part of web CNL cloud frontend located Computers Networks Laboratory at Technical University of Kosice. Paper explains streaming concepts in order to explain the need for monitoring and proactive solution of possible states where required services could not be provided. The main objective of this work is to design and implement the system that would display crucial streaming statistics of Videoserver using existing market available monitoring tools. Based on analysis appropriate monitoring tools are presented as part of this paper. The result is a website integrated as part of CNL frontend services and a mobile application for Android operating system, which also displays various realtime statistics and graphs of Videoserver usage. These solution is designed to help administrators with Videoserver management.

**Keywords** — Cacti, Monitorix, streaming, Videoserver, Wowza Media Server

## I. INTRODUCTION

As many internet usage statistics shows streaming is becoming widespread technology, offered services to desktop and mobile users. People like watching videos and clips online, live streaming from sport exhibitions, etc. In today's let say Internet age there are many people who often like to share some interesting videos that would like to share with their friends, colleagues or even public. Fact that i.e. YouTube has billions of videos watched daily shows how popular is this technology worldwide. Various software companies have realized this fact and put increasing emphasis on software development in this area. There is variety of solutions for streaming, redistribution and delivery/play on end device available. It is clear that streaming technology has big potential for future growth. One part of the game is to offer good level of monitoring services in order to allow simpler administration and to predict potential failures or bottlenecks within implemented topology.

## II. ANALYSIS

### A. The Videoserver platform

The Videoserver is the platform hosted on several virtual machines with one of them set as primary during normal operation state (decision in based on several system pre-defined thresholds) and serves as storage and video delivery service. This technology uses flash based streaming server Wowza which includes codecs, transcoding tools and broadcasting extensions to ensure high-quality streaming of digital video over IP networks. The Videoserver use metadata to fully identify object type of all video clips stored in database for efficient distribution. [1]

### B. Monitoring tool - Cacti

One of the best known monitoring tool on a market is Cacti [2], an open source, web-based network graphing and monitoring software designed to use powerful RRDTool's data storage



and graphing functionality. Cacti provides a fast poller, advanced graph templating, multiple data acquisition methods, and user management features out of the box. All of this is wrapped in an intuitive, easy to use interface that makes sense for LAN-sized installations up to complex networks with hundreds of devices. On Cacti authors' website it is possible find full documentation, installation instruction, usage instructions and lots of useful troubleshooting hints. This monitoring tool is available for Windows and Linux, in this case Linux based packages were installed on monitoring server. For example Cacti can monitor: running processes, CPU usage, system load average, etc. As Cacti provides modular platform with open API, authors created module for Wowza server monitoring.

### C. Monitoring tool - Monitorix

Monitorix is a free, open source, easy to use monitoring tool designed to monitor many services and system resources. For example Monitorix can monitor: CPU usage, connection counts, running processes, system load average and many other parameters. It has been created to be used under production Linux/UNIX servers, but thanks to its simplicity and small size can be used on embedded devices as well. Monitorix is available for free on author website where it is possible to find complete documentation of this monitoring tool, installation instruction and many other useful hints. [3]

## III. MONITORING SYSTEM

Monitoring system implemented to the Videosever platform is composed of several components. For monitoring purposes Cacti and Monitorix presented in previous chapter was selected, Drupal module was created to integrate statistics into website portal and mobile application was created for mobile access. As shown on figure (Fig.1) monitoring system monitors stream to end users (U) and monitors selected Videosever parameters. Cacti and Monitorix provide graphs using Drupal module to the internal website and mobile application on administrators' device. Administrator can log in using same username and password to website or mobile application and access statistics and graphs of the Videosever.

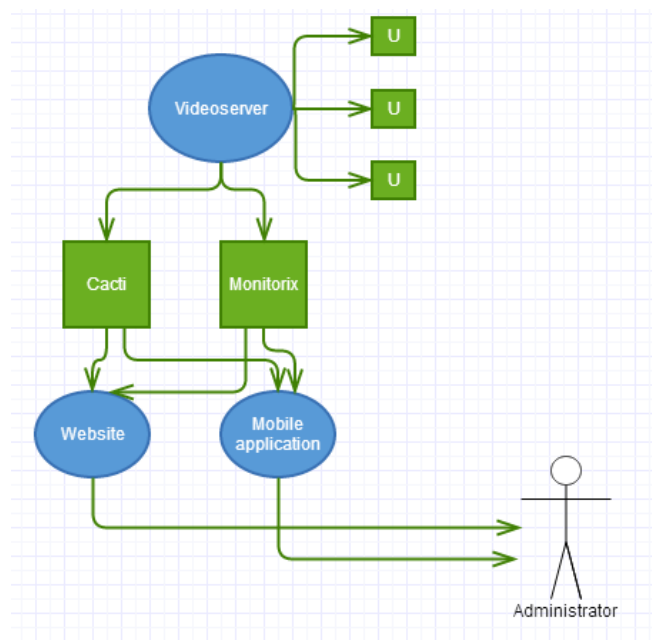


Fig. 1 Diagram of monitoring system

## IV. CONFIGURATION

### A. Cacti modifications

Basic configuration of Cacti is a quite simple task. As mentioned above, Cacti module for Wowza streaming server was created and implemented into monitoring portal. Due to extended need for cron services, existing poller.php file was altered and added to cacti system crontab.

```
*/5 * * * * cactiuser php /var/www/html/cacti/poller.php >
/dev/null 2>&1
```

Several Wowza streaming server system variables are monitored by Cacti module, i.e. number of connected users, number of users per event id, overall bandwidth or per event, number of multimedia resources mounted, etc. Basic Cacti monitoring is focused on primary streaming server and its linux based operating system parameters like CPU load, system load, IO operations, processes, network parameters, etc. In combination it provides quite good overview for system administrator regarding to state of streaming server.

### B. Monitorix modifications

Monitorix in basic configuration allows to measure Wowza streaming server service parameters. Only few additional modifications were required to measure current setup of the Wowza streaming platform. Configuration of Monitorix is primarily done within monitorix.cgi file. This file contains all configuration information and calls for external modules in case of need. In compare to Cacti it is possible to configure graph zoom, to set theme color or to configure measured scale. In this case it was necessary to do few modifications within RRD tool subsystem, system.rrd for advanced system load average and usage, and wowza.rrd for Wowza media server. Adding of Monitorix to platform serves as complementary source of information for the Videoserver platform administrator.

## V. SOLUTION

### A. Monitoring frontend

For access to the Videoserver platform results CMS Drupal module was created and added to monitoring platform (Fig. 2). Module supports Drupal in version 7, it is not depended on any other module and is theme independent. Created website is divided into several parts, which will be accessible through the main menu. The main menu includes project introduction, streaming statistics of Videoserver, memory statistics, processes list and its memory requirements, system load statistics, and other statistics of Videoserver focused primarily on Wowza streaming server. [4]

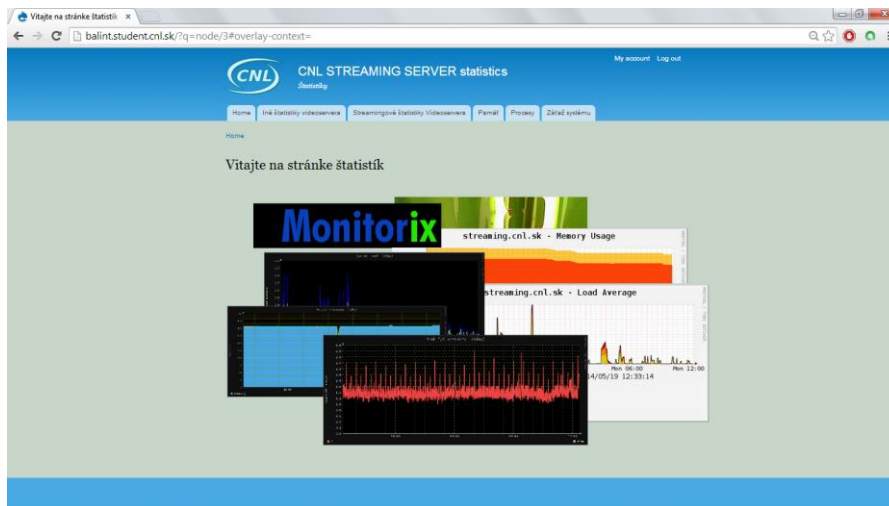


Fig. 2 Landing page of website prior to integration as module to existing Videoserver platform

### B. Mobile application for OS Android

As extension to Drupal module mobile application was introduced. It provides same amount of information as monitoring portal but is simpler to use on mobile devices. This application is designed to run on Android mobile version 2.2+ devices. Mobile application asks for login information (username and password), because some graphs and servers configuration could be understandable as private. Those information is not intended for the anonymous user, but only for the administrator.

Mobile application has simple functionality and design. It provides only authentication features and communication interface with server based API. In mobile application there is need to click on refresh button at the top for recent Videoserver information.

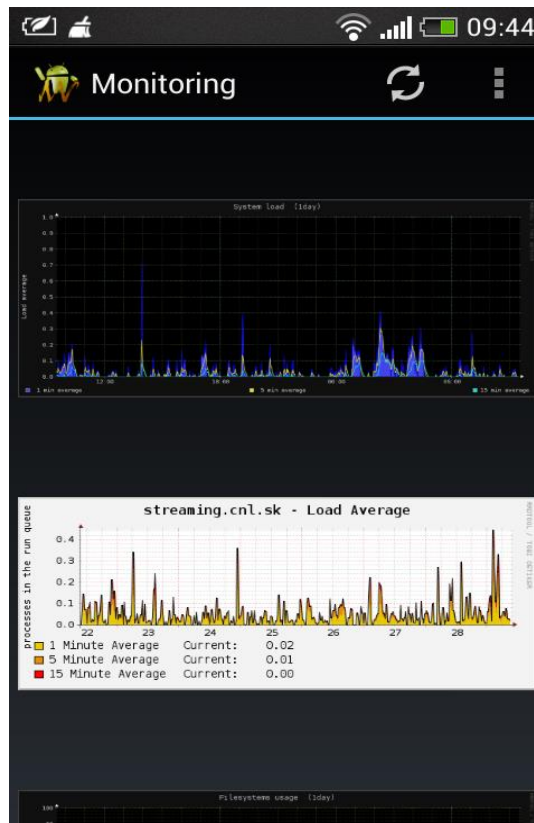


Fig. 3 Mobile application for Android OS

## VI. CONCLUSION

Papers presents implementation of two monitoring tools Cacti and Monitorix into the Videoserver streaming platform. Those monitoring tools were extended by new modules for monitoring of Wowza streaming server. In order to access statistics and graphs a Drupal module was created and implemented into CNL web frontend. For better mobility an Android application was created and tested. This work is a primary step for pro-active system module which is under development within Computer Networks Laboratory at Faculty of electrical engineering and informatics.

## ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia / Project is cofinanced from EU funds. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] Technopedia- definitions [online] 2013 [cit. 2014-10-06]  
Available on the internet: <<http://www.techopedia.com/definition/23872/video-server>>.
- [2] Cacti docs [online] 2013 [cit. 2014-10-06] Available on the internet: <<http://docs.cacti.net/>>.
- [3] Monitorix documentation [online] 2014 [cit. 2014-10-06]  
Available on the internet: <<http://www.monitorix.org/documentation.html#1>>.
- [4] Application for monitoring of streaming server. Filip Balint. Bachelor thesis. 2014. Technical university of Kosice.

# Real time object tracking and streaming Android application

<sup>1</sup>Richard LIPKA, <sup>2</sup>František JAKAB, <sup>3</sup>Dávid CYMBALÁK

<sup>1,2,3</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>richard.lipka@student.tuke.sk, <sup>2</sup>frantisek.jakab@tuke.sk, <sup>3</sup>david.cymbalak@cni.sk

**Abstract** — The main goal of this article is implementations of object tracking and real time streaming application for mobile platforms. In proposed solution user can choose an object on video sequence. When the user marks object in video, real-time tracking will start. The developer should select an effective library and algorithm, because mobile device has limited hardware. The algorithm evaluates every scene from the camera and compute information about object on the scene. This information is sent to the server. Video from the camera is sent via FTP protocol to the external server.

**Keywords** — Object, tracking, streaming, computer, vision, Android, application

## I. INTRODUCTION

In the modern applications as the autonomous systems, factory production and other sector use Computer Vision. Computer Vision is used all around us for example face detection in digital camera, red eye elimination, picture transformation and other.

Object tracking is part of Computer Vision. The user chooses an object on the picture from the camera, which he wants to track. Object tracking application finds a selected object on video sequence.

Streaming is a process where multimedia data are transforming continuously to another site.

## II. ANALYSIS

The goal of this work is to develop system consist of Android device and server. Android application is a basic part of the system. Android application consists of two parts. First part of this work is real time object tracking algorithm. Object tracking is a part of Computer Vision.

Computer vision [1] is a very important for object tracking. CV transforms data from picture or video to computer representation. The main goal is to create a vision similar to human vision. It is hard because the computer does not know to learn. Transformed data are represented by a number in a grid, where number represents value for the pixel Fig. 1. Some library can be used in CV application. CV libraries are libraries, which contents basic function and algorithm to work with pixels. FastCV and OpenCV are native libraries and BoofCV is Java library. Object tracking is a problem, which it is still in development. In this moment exists a lot of algorithm for object tracking for example: SURF, CAMSIFT, SAD, TLD and other. Many algorithms have problems with tracking when object changes size, changes position in the picture, rotated and deformed. Lighting conditions are very important for tracking on the scene. On google play store exists computer vision application. They are augmented reality application iOnRoad Augmented Driving, OpenCV FaceDetection, BoofCV demonstration, OpenTLD are demonstration computer vision application.

Second part of this work is real time video streaming. Video from device camera is streaming for third part u. exists a several method, how can be a video streaming implemented. For example video is streaming to server or image sequences are sending to server. Another way is to create a server at the mobile device. User can be connected with this stream via Media player. Media player is connected to stream via network using IP address and protocol. The server on

Android device can be created by using some library. Libstreaming and NanoHttpd is Java library to create a server on an Android device

We used native library OpenCV in this work, because the mobile device has limited hardware. It is very important to select a good object tracking algorithm, because the

application is real time. Method for video stream is important, because object tracking and video streaming are parallel process. The main goal is to choose a good and effective algorithm and method for video streaming.

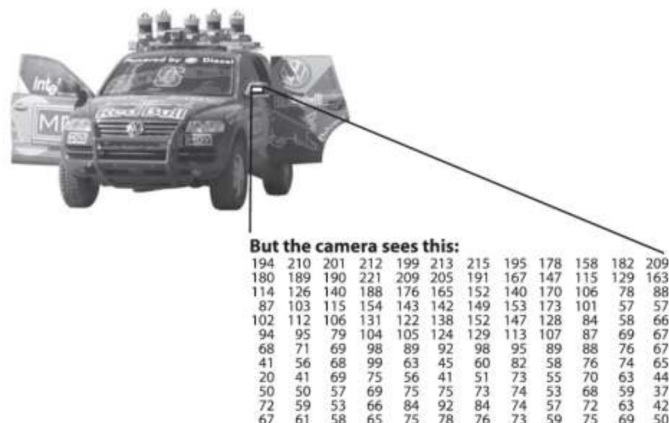


Fig. 1 Pixel data grid

### III. DESIGN AND SOLUTION

#### A. Application requirements

The main problem is how to do an effective real time object tracking and video streaming application. The ambition is implement a parallel application, where the streaming does not restrict object tracking.

This work must meet these requirements:

- When application was started, user should select object on the scene (Slide with finger on screen).
- Tracking object on video sequence, which was selected by user.
- If internet connection is enabled.
  - Every picture from video sequence is uploaded on the server. (Via FTP protocol)
  - Calculate information about an object in the scene. Information about the object are the width and height in pixel, object position and object representation in the picture.
  - The data obtained are sending in a Mysql database. These data can be used in other applications.

#### B. Design object tracking and video streaming application

It has been said that design application is very important. The application consists of two parts. Each part is implemented in own thread. The first thread is Upload thread. This thread sends pictures and information about scene to the server. The second thread is object tracking thread. Application consists of two threads, because object tracking and image uploading at server cannot influence. Application describe in follow steps:

- Select object.
- Calculate information about object.
- Send image and information about scene in upload thread from tracking thread.
- Create object from image and image data.
- Object with image and information about the image is inserting in stack. (Type of stack is FIFO).
- First image, which was inserted in stack is sent to server and information about object is inserting in external database. The Fig. 2 shows the parallel between tracking and streaming.

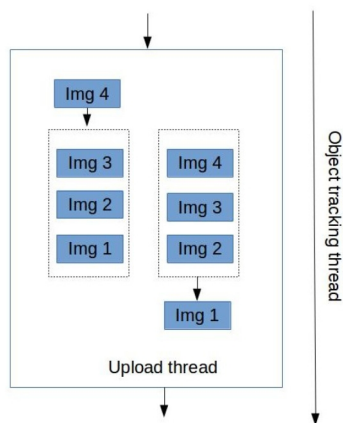


Fig. 2 Application design.

#### IV. IMPLEMENTATION OF APPLICATION

**Send data on server.** Each picture from the camera is uploading at server via FTP. SimpleFTP is a library, which was used for implementation FTP client on Android device. To calculate information about scene are inserting at external Mysql database via HTTP request (method POST). If PHP script gets requirement then insert data in the database.

**Object tracking.** OpenCV is the library, which was selected for implementation object tracking. It is robust, native library which includes more than 2000 computer vision methods. The implanted tracking algorithm is a modification of TLD algorithm.[2] The TLD is the algorithm, which was developed by Zdenek Kalal. TLD is a short cut which means tracking, learning and detection. The figure show how to TLD work. Detector compares model with actual picture from the camera. Model consists of a few patterns which represent objects in the scene. The model grows proportionally to representation object in the scene. Model with pattern represents learning part. The Fig. 3 shows that a tracking is recursive. Model consists of two patterns in this work. First patterns represent first selection from image. This pattern is stable. The second pattern is changing one and it represents learning part.

**TLD implementation.** Tld is Algorithm using optical flow and good features to track. As a first algorithm calculate good feature to track from picture. Shi-Tomasi corner detector calculate good feature.[3] Delete points from list of values, which not fall into an area of interest. Points in lists after deleting represent object. Good feature to track must be calculated from every picture.

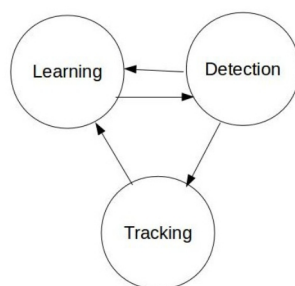


Fig. 3 TLD algorithm.

Optical flow [4] is the kind of method how we can calculate motion object between two neighboring frames. 2d vector represent pixel motion between first and second frame. The Fig. 4 show optical flow. First square represents position on first frame and the second one on next frame. Arrow symbolize 2d motion vector. Optical flow must satisfy the following condition. The pixel intensity cannot change between two neighboring frames. Neighboring Pixel has similar motion. Lucas-k

Kanade method was used in this work for calculating optical flow. This method takes 3x3 pixel around the point. All of nine points must have same motion.

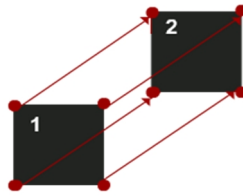


Fig. 4 Optical flow.

#### ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia/This project is being co-financed by the European Union. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### CONCLUSION

The main goal of this work was design and implementation of real time object tracking and streaming application. Application should be designed for Android devices. Object tracking is a part of computer vision.

The key was selected suitable object tracking algorithm and computer vision library. Application design begins, when it was selected. Design is important, because of the two tasks running at the same time in this application and mobile device has limited hardware. The main task is that streaming must minimal limiting object tracking.

Application was successfully implemented. This is the first mobile application, which tracks object, streaming video to the server and sending data about the object in external database.

#### REFERENCES

- [1] BRADSKI, Gary; KAEHLER, Adrian. Learning OpenCV: Computer vision with the OpenCV library. O'reilly, 2008. ISBN 978-0-596-51613-0
- [2] NEBEHAY, Georg. Robust Object Tracking Based on Tracking-Learning- Detection. Vienna: Faculty of Informatics, Vienna University of Technology, 20120.
- [3] SHI, Jianbo; TOMASI, Carlo. Good features to track. In: Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on. IEEE, 1994. p. 593-600.
- [4] AIRES, Kelson RT; SANTANA, Andre M.; MEDEIROS, Adelardo AD. Optical flow using color information: preliminary results. In: Proceedings of the 2008 ACM symposium on Applied computing. ACM, 2008. p. 1607-1611. ISBN 978-1-59593-753-7.

# Conditional video streaming from mobile devices in multi-cameras environment

<sup>1</sup>Lukáš HOLUBEC, <sup>2</sup>František JAKAB and <sup>3</sup>Dávid CYMBALÁK

<sup>1,2,3</sup> Department of Computers and informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>lukas.holubec@student.tuke.sk, <sup>2</sup>frantisek.jakab@tuke.sk, <sup>3</sup>david.cymbalak@tuke.sk

**Abstract** — The target of this paper is to design and create an applications for receiving the group of video streams distributed over a network from mobile devices. The application will identifies the location of known object with use of tracking algorithm in the image in real time. The developed program evaluates on the basis of an evaluation mechanism position with the entire group of video stream. These reviews are regularly stored into the database and prototype user interface is use to display all current rating of streams videos and to reproduce best rated video stream.

**Keywords** — SURF, OpenCV, tracking algorithm, computer vision, recognition image, mobile devices

## I. INTRODUCTION

The interest for computer vision technology is rapidly growing domain today. These systems are becoming part of everyday life. Computer visions systems are part of the systems quality control of products, security systems, driving assistance systems, etc. These systems are highly specialized and comprehensive solutions designed to meet clearly defined task. Such solutions are usually adapted to the individual circumstances of each customer. They require specific expensive hardware and software solutions.

There is a challenge to create similar solution but cheaper and based on the mobile device what is using more and more in everyday life. This solution will could use everybody according to their needs.

This work deals with the design, create application and experimental verification solutions for receive and processing of video from multiple mobile devices with the use of tracking algorithms.

## II. GOALS

The aim of this work is the design and implementation application for income group video streams what it distributed over a computer network and it produced by mobile devices. The application identified through tracking algorithm position of the pattern object within the image in real time. Application evaluates the image quality on the object through an evaluation mechanism, which is design in this work and evaluation of frame will be entered in the database.

## III. ANALYSIS

Transmission of audiovisual content over a network has increasingly larger share of the data transmitted over the Internet. According to the Cisco Visual Networking Index Forecast will be in 2016 transmitted by the Internet 1.2 million minutes of video content per second. This trend requires the development new solutions for transmission and processing of video content transmit over a network.

Video content can be transmitted over a computer network through two technologies. The first technology is a technology unicast. This technology is suitable for transmitting video stream for a small group of receivers or for transmission in the local computer network because the stream server sends one copy of video stream each receiver. The second technology is a technology multicast. This technology is an appropriate solution for sending video stream in



large group of receivers. Stream server sends only one copy of the video stream, which it replicates on the transition nodes computer networks. For the purpose of this work is more appropriate technology unicast because stream server sends video stream only for one receiver.

Standard used to transfer video content is next important element. Standards are divided by several criteria. Transmission standards are divided on loss and lossless standards. Lossless codec that uses lossless video standard do not change video quality but the resulting compression ratio is only 1:2. Loss codec that uses loss video standard reduce video quality but loss codec have a compression ratio up to 1:100. For using a computer network on transmit video content, it is necessary to use advanced codecs, which are educe from loss video standards. H.264 is the most used standard for video transmission. H.264 has 80 % market share. This standard is licensed but its use is free. VP8 is the most used unlicensed video standard. H.264 and VP8 are old video standards; we need better video standards for future. H.265 and VP9 are successor's old standards H.264 and VP8.

Computer vision is discipline of computer science. Computer vision includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data from the real world in order to produce numerical or symbolic information [1]. For this work we are interested for domain of tracking algorithms. Domain tracking algorithms and algorithms of recognition is young domain. The computers are in domain object recognition on same level than two years old children today [2].

Object can be recognized by several methods by measurable symptoms, structures elements or models. Object recognition by measurable symptom of the image looking for measurable parameter than brightness or color. Object recognition by model uses 2D, 3D models or a precise mathematical representation of the object. This method is looking for projections of the models in the scene. Object recognition by structural comparison compares the base structure (points, lines, edges) of the object and scene. Tracking algorithms have not good results always because object recognition is very computing process and tracking algorithms are sensitive on quality image and pattern.

#### IV. SOLUTION AND RESULTS

##### A. The design of solution

This work does not design or create application for recording and streaming video by computer network. The work uses a free available application on recording and transmission of video information distributed by google apps. The mobile application was chosen according to acquired knowledge about the transmission of video by computer network.

The solution is created using modified software pattern model, view, and controller in Fig. 1. View module generates a graphical user interface to communicate with the user of application. Unit model is a database that is used for communication applications with other applications. The database contains output data applications, which are individual identification video streams and rating of individual video streams. Controller module is responsible for managing and implementing the various activities of the program. Controller module has three parts: control part, identify part, rating part in Fig. 2. Control part manages the activities of the program. Identify part searches using tracking algorithm model object in the scene. Identify part produces data for rating part application. Rating part rate position of identify model in scene. Data form rating part is stored in database for other application.

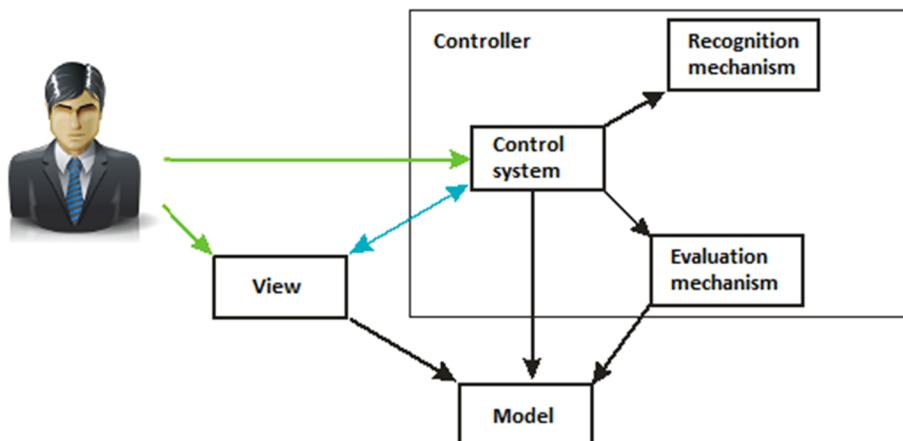


Fig. 1 Software pattern model, view, controller

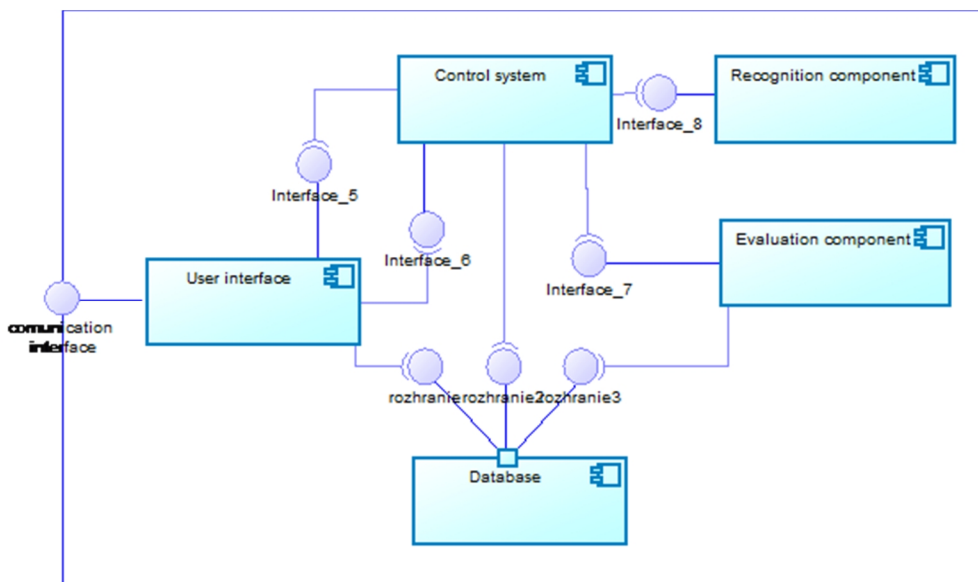


Fig. 2 UML component diagram

### B. Pattern recognition

In the work was used recognition algorithm SURF (Speeded Up Robust Features). SURF is a robust local feature detector that can be used in computer vision tasks like object recognition or 3D reconstruction [3]. SURF is inspired by the SIFT descriptor. SURF is designed to recognize one object in the scene by model. Implementation of recognition at work consists of 5 points:

1. Detection of unique key points in the model and unique identification this points.
2. Detection of unique key points in the scene and unique identification this points.
3. Looking for similar key points in model and scene by usage method brute force.
4. Control and elimination of faults, a pair of points (criterion is the uniqueness of these pairs).
5. Limiting the range of the recognition model in the scene with four points.

### C. Rating video stream

Image rating has two components. The first component is the degree similarity of the object model with object recognition. During the recognition phase, each pair of detection points assigned a value. This value reflects the match point in model and recognized point in the scene. After the filtering process I can determine the similarity of the model and recognized object by detected pairs of points. Recognized pairs of points have a similar degree of similarity as a pairs who have not been recognized. Degree of similarity the model and the detected object are measured by use statistical method (median).

The second component expresses ideality location object in the frame. We evaluate the position center of the object to the center of the frame. Center of the object is the intersection of the diagonals of the area in which the object is detected Fig. 3. Rating is the sum of x and y components (1) (2). Components reflect the ratio of the distance from the center of the object to boundaries of permissible areas and distance from the center of the frame to boundaries of the permissible area (3). Position boundaries of the permissible area from the edge frame are equal distance center of object to outermost point of object in this direction Fig. 4.

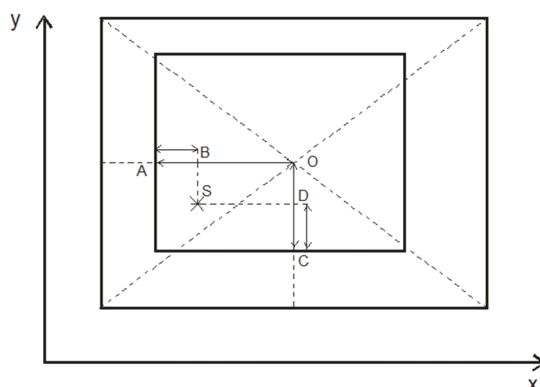


Fig. 3 Determine the components of x and y

$$D = \begin{vmatrix} x_B - x_A & x_P - x_Q \\ y_B - y_A & y_P - y_Q \end{vmatrix} \quad D_t = \begin{vmatrix} x_P - x_A & x_P - x_Q \\ y_P - y_A & y_P - y_Q \end{vmatrix} \quad t = \frac{D_t}{D} \quad (1)$$

$$x = x_A + t(x_B - x_A) \quad y = y_A + t(y_B - y_A) \quad (2)$$

$$x = \frac{|AB|}{|AO|} \quad y = \frac{|CD|}{|CO|} \quad (3)$$

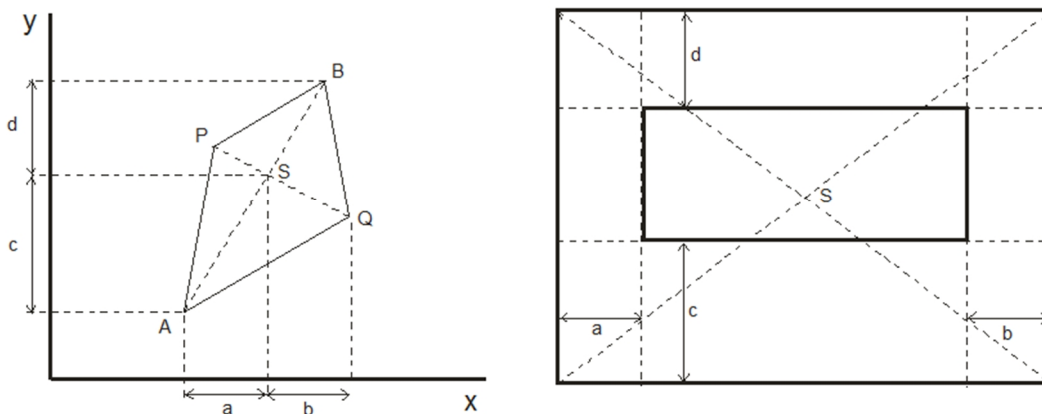


Fig. 4 Determine the boundaries of the area

#### V. CONCLUSION

The aim of this work was design and implement application based on an analysis of available tracking algorithms and with using free available mobile application for streaming video from mobile devices. Applications receiving parallel multiple streams video in which recognizes the model object. App rating each video stream by quality frame on the object. The application stores the evaluation of each video stream to the database (for other programs).

The benefit of this solution is the use of tracking algorithms, with more and cheaper mobile devices and standard computers.

#### ACKNOWLEDGMENT

Paper/poster/article is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia/This project is being co-financed by the European Union. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] R Klette, Concise Computer Vision: an introduction into theory and algorithmus, in *Springer*, London, 2014, ISBN: 978-1-4471-6320-6
- [2] R. Szeliski, Computer Vision: Algorithms and Applications, in *Springer*, London, 19. Oct. 2010, ISBN: 1848829345, pp. 9–20
- [3] L. D. Baggio, S. Emami, M. D. Escrivá, K. Levgen, Mastering OpenCV with Practical Computer Vision Projects, In Packt Publishing Ltd., Birningham, Dec. 2012, ISBN: 978-1-84951-782-9, pp. 95-109

# Eliminácia vplyvu prechodných dejov na elektrickú sieť

Vladimír IMRICH, Dušan MEDVEĎ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
univerzita v Košiciach, Slovenská republika

vladimir.imrich@yahoo.com, dusan.medved@tuke.sk

**Abstrakt** — príspevok sa zaoberá analýzou prechodných dejov spôsobených prevádzkou elektrickej oblúkovej pece v sieti 22 kV, popisuje príčiny ich vzniku a navrhuje možnosti ich eliminácie s ohľadom na splnenie kvality elektrickej energie. Zameriava sa najmä na vyhodnotenie krátkodobej miery blikania (flickra) a využitie zariadenia STATCOM pri jej eliminácii. Analýza i následná eliminácia je realizovaná simuláciami pomocou nástroja Matlab/SimPowerSystems. Získané výsledky sú vyhodnotené v súlade s platnými európskymi normami a legislatívou SR.

**Kľúčové slová** — prechodné deje, distribučná sústava, elektrická oblúčová pec, krátkodobá miera blikania, flicker, STATCOM

## I. ÚVOD

Určité druhy prechodných dejov neovplyvňujú iba chod elektrizačnej sústavy (ES), ale svojimi prejavmi vplývajú priamo na človeka. Takýmto prejavom prechodných dejov je miera blikania (flicker), ktorá dynamickými zmenami napätia spôsobuje blikanie svetelných zdrojov, ktoré nepriaznivo vplývajú na pohodu osôb. Silný zdroj takýchto prechodných dejov predstavuje prevádzka elektrickej oblúkovej pece (EOP), pretože dovolená miera blikania nebýva dodržaná skôr ako ostatné podmienky povoľujúce prevádzku v danom bode siete. Problematike pridáva na aktuálnosti i fakt, že rozvoj výkonových polovodičových prvkov v poslednom období umožňuje v sústave nasadzovanie zariadení FACTS (flexible AC transmission systems). Tie svojimi vlastnosťami spĺňajú predpoklady spoľahlivej kompenzácie prechodných dejov spôsobených EOP bez rozsiahlych zásahov do usporiadania siete.

Problematika bola skúmaná na príklade fiktívneho pripojenia elektrickej oblúkovej pece o zdanlivom výkone 2 MVA do distribučnej siete 22 kV na vedenie V-281 pripojené do elektrickej stanice Bardejov.

## II. PRECHODNÉ JAVY SPÔSOBENÉ PREVÁDZKOU EOP

V článku boli skúmané prechodné deje vznikajúce pri prevádzke EOP, ktoré patria do kategórie elektromagnetických prechodných dejov. EOP svojou prevádzkou predstavuje nepravidelné prúdové zaťaženie siete spôsobujúce na impedanciách siete poklesy napätia, ktoré majú za následok kolísanie napätia v sústave. Podľa [1] je spomínané kolísanie spôsobené 2 typmi prúdového zaťaženia vznikajúceho v EOP označenými ako prvý a druhý typ zmien.

### A. Prvý typ zmien prúdového zaťaženia

Vyskytuje sa najmä v prvých 30 minútach tavby s frekvenciou výskytu 0,5- až 1-krát za sekundu a predstavuje skratový prúd pri zapálení oblúka medzi elektródou a vsádzkou. V obvode EOP sa jeho výskyt neustále opakuje dokiaľ je vsádzka studená. Ďalšia možnosť vzniku zmeny tohto typu je pri náhodnom kontakte vsádzky s elektródou (napr. posunutím vsádzky počas procesu tavby).

### B. Druhý typ zmien prúdového zaťaženia

Predstavuje zmenu intenzity prúdu v rozsahu  $\pm 15$  až  $\pm 50$  % menovitej hodnoty s frekvenciou výskytu 2- až 20-krát za sekundu. „Zdá sa, že tieto kolísania prúdu nepodliehajú jednoznačnej zákonitosti, pretože sa stále menia ako vo svojej amplitúde, tak aj početnosti.“ [1] Prúdové zmeny a nimi spôsobené kolísanie má 3 príčiny:

- Zmeny intenzity prúdu bez výraznejších zmien dĺžky elektrického oblúka pri prechode prúdu z jednej polvlny do druhej. Vplyvom teploty sa materiál vystavený oblúku roztaví a sčasti vyparí, pričom silným miestnym prehrievaním môže vyparovanie prísad v oceli prebiehať explozívne. Vzniknuté pary stúpajú a v okolí oblúka sa vplyvom vysokých teplôt ionizujú, čím zvyšujú koncentráciu voľných nosičov a umožňujú oblúkom preniesť vyšší prúd.
- Kolísanie spôsobené preskokom oblúka z jedného vyčnievajúceho kusu vsádzky na druhý. K preskoku dochádza s periódou 0,1 až 0,4 sekundy (2 až 10 preskokov za sekundu), pričom preskok nastane zakaždým, keď sa vyčnievajúci kus vsádzky vplyvom elektrického oblúka zaoblí. Zmeny amplitúdy prúdu spôsobené preskokmi oblúka sú hlavnou príčinou blikania svetla (flicker efekt).
- Predlžovanie oblúka z dôvodu jeho slučkového pohybu. Vychýlenie oblúka je zapríčinené zosilnením magnetického poľa susednej fázy, ktorou z určitej vyššie popísanej príčiny začal tiecť väčší prúd.

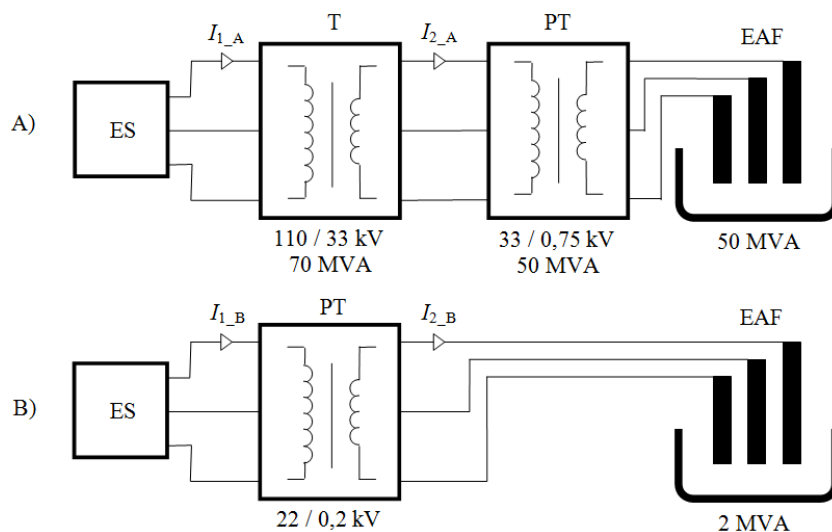
### III. POUŽITÉ METÓDY PRE ANALÝZU PRECHODNÝCH DEJOV

Eliminácia, ako aj samotná simulácia nesymetrickej záťaže v podobe EOP bola vykonaná prostredníctvom programu Matlab v grafickom prostredí Simulink. Pre modelovanie prvkov silových elektrických systémov bol využitý nástroj SimPowerSystems rozširujúci Simulink o knižnicu týchto prvkov. Pomocou spomínaných knižníc boli vytvorené modely nadradenej siete, transformátorov, vonkajších a káblových vedení, kompenzačnej tlmivky a tiež modely symetrických záťaží na strane nízkeho napätia.

#### A. Model elektrickej oblúkovej pece

Model oblúkovej pece pozostáva z krátkej siete a elektród. Tieto prvky s ich elektrickými parametrami sú v modeli nahradené ekvivalentnou hodnotou prúdu, aký by nimi bol odoberaný zo svoriek sekundárnej strany pecného transformátora. V tomto prípade je spomínaný prúd určený hodnotami prúdov nameranými pri skutočnej prevádzke elektrickej oblúkovej pece výkonu 50 MVA, pripojenej k elektrickej sieti cez pecný transformátor 110/33 kV. Prúdy boli merané na sekundárnej strane transformátora, pričom hodnoty boli počas 1-minútového merania odčítavané každú polperiódu. Tieto hodnoty prúdov boli prispôbené tak, aby zodpovedali prevádzke oblúkovej pece s výkonom 2 MVA pripojenej cez pecný transformátor 22/0,2 kV.

Pre odvodenie ekvivalentného správania sa oblúkovej pece v sieti 22 kV (Obr. 1 B) sa vychádzalo z predpokladu, že zo siete odoberá rovnaký prúd  $I_2$  v pomere k nominálnemu, ako oblúková pec (A) napájaná zo siete 110 kV.



Obr. 1 Princiálne znázornenie EOP s pecným transformátorom a ich hlavnými parametrami  
 Rovnosť týchto pomerov sa zapíše ako:

$$\frac{I_{2\_A}}{I_{2n\_A}} = \frac{I_{2\_B}}{I_{2n\_B}} \quad (1),$$

odkiaľ úpravou získame vzťah:

$$I_{2\_B} = \frac{I_{2\_A}}{I_{2n\_A}} \cdot I_{2n\_B} = \frac{I_{2n\_B}}{I_{2n\_A}} \cdot I_{2\_A} \quad (2)$$

Keďže:

$$S_n = \sqrt{3} \cdot U_{2n} \cdot I_{2n} \quad (3)$$

potom:

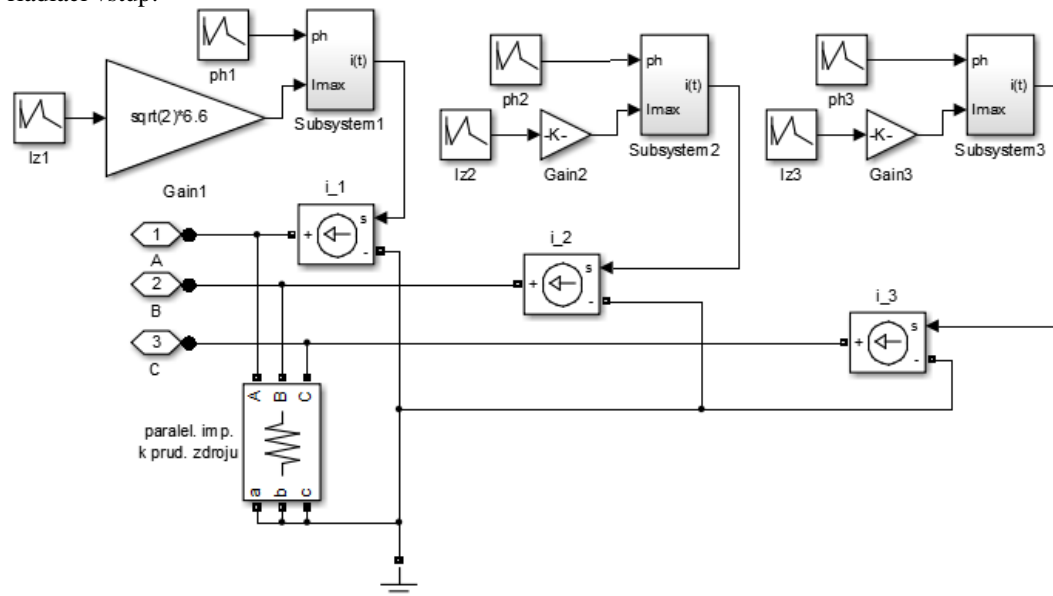
$$I_{2n\_A} = \frac{S_{n\_A}}{\sqrt{3} \cdot U_{2n\_A}} \quad I_{2n\_B} = \frac{S_{n\_B}}{\sqrt{3} \cdot U_{2n\_B}} \quad (4)$$

Dosadením výrazu (4) do (2) získame vzťah:

$$I_{2\_B} = \frac{S_{n\_B} \cdot U_{2n\_A}}{S_{n\_A} \cdot U_{2n\_B}} \cdot I_{2\_A} = \frac{2 \cdot 10^6 \cdot 33 \cdot 10^3}{50 \cdot 10^6 \cdot 0,2 \cdot 10^3} \cdot I_{2\_A} = 6,6 \cdot I_{2\_A} \quad (5)$$

pričom výsledok predstavuje konštantu, ktorou sa musí prúd získaný meraním na 50 MVA oblúkovej peči vynásobiť, aby zodpovedal zaťaženiu 2 MVA pece.

Základ pre model oblúkovej pece preto tvorí blok **Controlled Current Source**, teda riadený prúdový zdroj (Obr. 2 i\_1 až i\_3), ktorého prúd pôsobí proti prúdom tečúcim v sieti. Riadený znamená, že hodnota prúdu sa môže počas behu simulácie meniť podľa hodnôt privedených na riadiaci vstup.

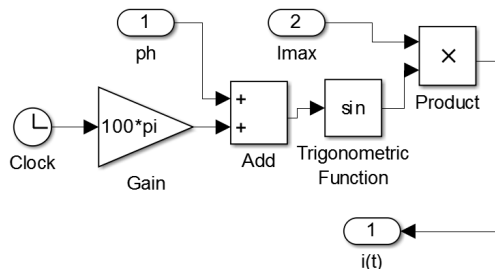


Obr. 2 Bloková schéma modelu oblúkovej pece

Keďže prúd nastavený v tomto zdroji má byť vlastnosťami zhodný s prúdom v ES, definuje sa riadiaci vstup nasledovne:

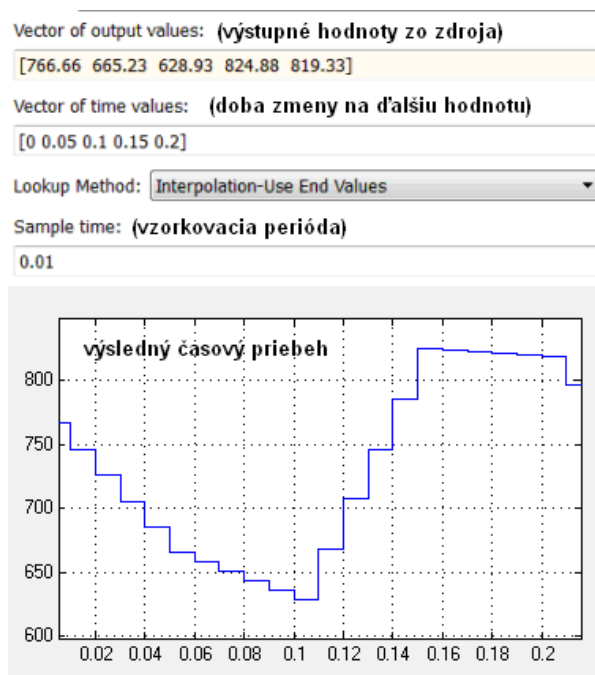
$$i_n = I_{mn} \cdot \sin(\omega \cdot t + \varphi_n), \text{ kde } n \text{ predstavuje označenie fázy} \quad (6)$$

Tieto vzťahy sú na schéme na Obr. 2 reprezentované vytvorenými blokmi Subsystem1, 2 a 3, ktorých vnútorné zapojenie je na Obr. 3. Vstupy do týchto blokov predstavujú hodnoty veličín  $I_{mn}$  a  $\varphi_n$ .



Obr. 3 Bloková schéma bloku pre vytvorenie sínusového signálu

Za zdroj signálu s hodnotami  $I_{mn}$  a  $\varphi_n$  bol zvolený blok **Repeating Sequence Interpolated**, ktorého činnosť je ilustrovaná na Obr. 4.



Obr. 4 Príklad činnosti bloku Repeating Sequence Interpolated

Signál z tohto bloku teda mení svoju hodnotu zakaždým po uplynutí vzorkovacej periódy, pričom žiadanú hodnotu nadobudne až v požadovanom časovom okamihu na dobu jednej vzorkovacej periódy. Postupné čiastkové zmeny medzi dvoma požadovanými hodnotami majú lineárny charakter. Frekvencia týchto prúdových zmien bola pre oblúkovú pec nastavená, vzhľadom k teoretickým poznatkom uvedeným v podkapitole II.B, a má hodnotu 20 Hz. K zmene odoberaného prúdu z oblúkovej pece teda dochádza plynule, pričom maximum prúdovej zmeny sa dosahuje s periódou 0,05 sekundy.

#### B. Model flickermetra

Model flickermetra nie je v prostredí SimPowerSystems vytvorený v samostatnom bloku, avšak je vytvorený medzi vzorovými príkladmi, ktoré obsahujú knižnice Matlabu. Model flickermetra zodpovedá štandardu IEC 6100-4-15 a jeho charakteristika je bližšie popísaná v [2]. Model umožňuje v nastaveniach zvoliť frekvenciu siete, ktorú ide analyzovať. V niektorých blokoch, ktoré obsahuje však napriek tomu ostáva nastavená predvolená frekvencia 60 Hz, ktorú je potrebné prepísať ručne.

Spomínaný model umožňuje meranie okamžitej miery blikania (IFS – instantaneous flicker sensation) iba v jednej fáze, čo pri vyhodnocovaní nesymetrickej záťaže v podobe oblúkovej pece, nepostačuje. Daný model bol z tohto dôvodu ďalej upravený tak, aby v určenom bode sústavy dokázal vyhodnotiť mieru blikania vo všetkých troch fázach.

Dovolené hodnoty okamžitej miery blikania nie sú uvedené v norme, preto bolo potrebné hodnoty získané flickermetrom ďalej vyhodnotiť na hodnotu krátkodobej miery blikania  $P_{st}$ . Krátkodobá miera blikania bola určená podľa [3].

#### C. Model zariadenia STATCOM

Knižnica prostredia SimPowerSystems obsahuje niekoľko blokov zariadení FACTS, avšak všetky sú určené pre simulácie s použitím fázorovej výpočtovej metódy, ktorej použitie na skúmaný problém nevyhovovalo požiadavkám. Ako model bolo preto, podobne ako model flickermetra, zvolené využitie zariadenia STATCOM namodelované v príklade. Jedná sa o diskretný model STATCOM-u s nominálnym napätím 25 kV a možnosťou regulácie výkonu v rozsahu  $\pm 3$  MVar v sústave s frekvenciou 60 Hz. Pre jeho použitie v skúmanej ES bolo potrebné zmeniť nastavenie frekvencie regulátorov na 50 Hz otvorením podblokov, ktoré obsahuje. Bližší popis modelu je popísaný v literatúre [4].

#### D. Simulácie siete

V simuláciách siete bola analyzovaná miera blikania (okamžitá a krátkodobá) pri rôznych opatreniach na jej zníženie. Čas simulácie bol vo všetkých prípadoch 60 sekúnd. Tento čas bol zvolený z dôvodu, že 10-minútové simulácie potrebujú pre určenie krátkodobej miery blikania podľa normy STN EN 50160, boli z hľadiska výpočtového času použitých počítačov, takmer nere realizovateľné. Flickermeter bol zapojený vždy v uzle pripojenia EOP k vedeniu V-281 predpokladajúc, že ak bola splnená kvalita elektriny v danom uzle, bola splnená tiež pre celý



zvyšok siete. Pre náročnosť výpočtu bol flickermeter umiestnený iba v spomínanom uzle a v ostatných uzloch bolo merané iba napätie. Dáta získané meraním boli exportované do súborov a spracované pomocou programu Microsoft Office Excel.

#### IV. VÝSLEDKY

Ako prvá bola realizovaná simulácia s kompenzáciou výlučne pomocou kompenzačnej tlmivky, pričom EOP bola pripojená do uzla siete AV. Z dôvodu nespĺnenia dovolenej krátkodobej miery blikania (Tab. 1) sa pristúpilo ku kompenzácii pomocou zariadenia STATCOM a pri neustále prekročenej hodnote  $P_{st}$  bolo, ako ďalšia možnosť eliminácie flickra, uvažované zvýšenie skratového výkonu v elektrickej stanici.

Tab. 1 Porovnanie hodnôt  $P_{st}$  pri použití jednotlivých úprav v sieti

Bod pripojenia / spôsob kompenzácie	$P_{st}$		
	fáza A	fáza B	fáza C
Uzol AV / Tlmivka (TL)	9,217	19,187	23,554
Uzol AV / Tlmivka + STATCOM (S)	4,372	9,384	12,611
Uzol AV / Tlmivka + $3 \cdot S_{k3}''$	9,080	19,051	23,417
Uzol AV / TL + S + $3 \cdot S_{k3}''$	4,340	9,289	12,512
Uzol AV / TL ( $3 \cdot L$ ) + S + $3 \cdot S_{k3}''$	3,878	5,088	5,946
El. stanica (22 kV) / TL + S <sup>(1)</sup>	1,204	1,427	1,332
El. stanica (22 kV) / TL + S <sup>(2)</sup>	5,123	11,301	14,772
El. stanica (22 kV) / TL + S + $3 \cdot S_{k3}''$	1,179	1,381	1,278
El. stanica (22 kV) / TL ( $3 \cdot L$ ) + S + $3 \cdot S_{k3}''$	1,177	1,360	1,238
El. stanica (110 kV) / TL + S	1,122	1,289	1,160
El. stanica (110 kV) / TL	1,115	1,289	1,150
El. stanica (110 kV) / TL - meranie na 22 kV	1,105	1,273	1,136
El. stanica (110 kV) / TL + $3 \cdot S_{k3}''$	1,098	1,260	1,121
El. stanica (110 kV) / TL + $3 \cdot S_{k3}''$ - mer. 22 kV	1,088	1,245	1,109

<sup>(1)</sup> - merané na strane pripojenia vývodu k sieti

<sup>(2)</sup> - merané na strane ukončenia vývodu pri EOP

TL - pripojená kompenzačná tlmivka

S - pripojené zariadenie STATCOM

$3 \cdot S_{k3}''$  - 3-násobné zvýšenie skratového výkonu v elektrickej stanici

TL ( $3 \cdot L$ ) - použitie tlmivky s 3-násobnou hodnotou indukčnosti

Porovnaním skratových výkonov elektrických staníc na východnom Slovensku je možné povedať, že ak by sme v danej elektrickej stanici Bardejov predpokladali inštaláciu transformátora 400 / 110 kV, zvýšil by sa jej skratový výkon približne trojnásobne, teda sme uvažovali  $S_{k3}'' = 3891$  GVA. Ani toto opatrenie neprinieslo žiadané zníženie hodnoty  $P_{st}$ , teda došlo následne k zvýšeniu indukčnosti kompenzačnej tlmivky na trojnásobok pôvodnej hodnoty.

Krátkodobá miera blikania však bola na úrovni 5-násobku dovolenej hodnoty, preto sa ako ďalšie riešenie uvažovalo pripojenie EOP s elektrickou stanicou pomocou vlastného VN vývodu. Týmto riešením došlo k výraznému zníženiu miery blikania v sieti, pričom je nutné poznamenať, že miera blikania poklesla vďaka zvýšeniu skratového výkonu v mieste merania flickra na začiatku vývodu a tlmiacim účinkom samotného vedenia napájajúceho EOP. Na konci spomínaného vedenia na strane EOP však miera blikania kvôli menšiemu podielu symetrickej záťaže narástla. Tento problém by sa však dal v pomocných inštaláciách EOP vyriešiť ich napojením z uzla AV vedenia V-281. Ani takéto riešenie kombinované spolu so všetkými predošlými nezabezpečilo splnenie legislatívou dovolenej hodnoty. Posledné realizované riešenie bolo pripojenie EOP s elektrickou stanicou pomocou samostatného vývodu zo strany 110 kV.

#### V. ZÁVER

Zo získaných výsledkov v kapitole IV. vyplýva, že pripojením oblúkovej pece do uzla siete AV, s použitím kompenzačnej tlmivky ako prvku zabezpečujúceho elimináciu prechodných dejov, dochádzalo k výraznému (24-násobnému) prekročeniu dovolenej krátkodobej miery blikania. Táto pomerne vysoká hodnota  $P_{st}$  bola spôsobená výberom vzorky nameraných prúdov



pri prevádzke pece, kde bolo zvolené meranie s čo najvyššími prúdovými špičkami a zároveň veľkým rozdielom medzi maximálnym a minimálnym odoberaným prúdom. Krátkodobá miera blikania teda bola vyhodnocovaná pre najnepriaznivejší prípad prevádzky EOP na začiatku tavby.

Ani jedno zo skúmaných riešení nedokázalo eliminovať krátkodobú mieru blikania zapríčinenú prechodnými javmi, na úroveň dovolenú v norme STN EN 50160. Uvažovaná 2 MVA oblúčková pec by teda na skúmanom vedení V-281 nemohla byť prevádzkovaná bez toho, aby nedošlo k nedodržaniu kvality napätia.

Z toho dôvodu je nutné dodať, že dlhodobá miera blikania, vyhodnocovaná v reálnych meraniach, by mala oproti získaným výsledkom nižšie hodnoty. Tento fakt je zapríčinený tým, že dlhodobá miera blikania je vyhodnocovaná z dvanástich 10-minútových meraní a musí byť dodržaná 95 % z doby jedného týždňa. Väčšina dejov spôsobujúcich veľké výkyvy odoberaných prúdov sa však odohráva iba v prvých 15 minútach každej tavby. V simuláciách bola ale miera blikania skúmaná iba po dobu jednej – najhoršej minúty prevádzky.

#### POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied grantom VEGA č. 1/0388/13.

#### REFERENCIE

- [1] HRADÍLEK, Z.: *Elektroenergetika distribučních a průmyslových zařízení*. Ostrava : Montanex, 2008. 364 s. ISBN 987-80-7225-291-6
- [2] WHITE, L. W.: *A discrete Matlab–Simulink flickermeter model for power quality studies*. [online]. [cit. 2014-03-12]. Dostupné na internete: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05378555>>
- [3] CABALA, L.: *Měření flikru s využitím virtuální instrumentace*. Bakalárska práca. Brno: VUT FEKT, 2010. 46 s.
- [4] GIROUX, P. a kol.: *Modeling and simulation of a distribution STATCOM using simulink's power system blockset*. [online]. [cit. 2014-03-18]. Dostupné na internete: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00975905>>

# Vplyv parametrov kontaktovania na tvar slučky kontaktovaného spoja

Igor VEHEC

<sup>1</sup> Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovensko

i.vehec@tuke.sk

**Abstrakt** — Kontrola tvaru slučky kontaktovaných spojov je jedným z dôležitých faktorov ovplyvňujúci kvalitu kontaktu. Špecifické aplikácie puzdrenia za stále zvyšujúcej sa integrácie môžu vyžadovať aj špecifické požiadavky na tvar spoja. V článku je načrtnutá jednoduchá metóda nastavenia parametrov kontaktovacieho zariadenia pre vytvorenie kontaktovaného spoja so štandardným trojuholníkovým tvarom slučky.

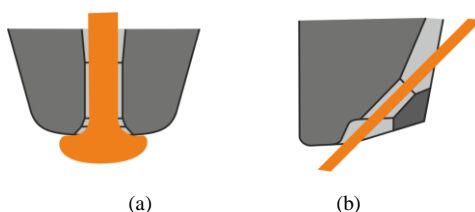
**Kľúčové slová** — kontaktované spoje, slučka kontaktovaného spoja

## I. ÚVOD

Kontaktované spoje majú v súčasnosti pri prepájaní elektronických komponentov na úrovni polovodičových čipov stále najväčšie zastúpenie. Moderným trendom v elektronike je zvyšovanie integrácie, ktoré je v oblasti kontaktovania spojené so znižovaním rozmerov kontaktovacích plôšok a rozstupov medzi nimi. Náročné aplikácie si často vyžadujú kontrolu tvaru slučky kontaktovaného spoja. Dôležitým problémom pri kontrole tvaru slučky kontaktovaného spoja je určiť trajektóriu kapiláry tak, aby vytvorené vodivé spojenie spĺňalo požiadavky ako z hľadiska puzdrenia, tak aj udržania výšky vytvoreného kontaktu. V súčasnosti prevláda v priemysle nastavenie parametrov kontaktovania systémom pokus-omyl, ktoré je časovo náročné a neefektívne. Aby bolo možné rýchlo a efektívne vytvoriť slučku s optimálnym profilom, je dôležité plne porozumieť dynamickým charakteristikám vytvárania slučky kontaktovaného spoja [1][2].

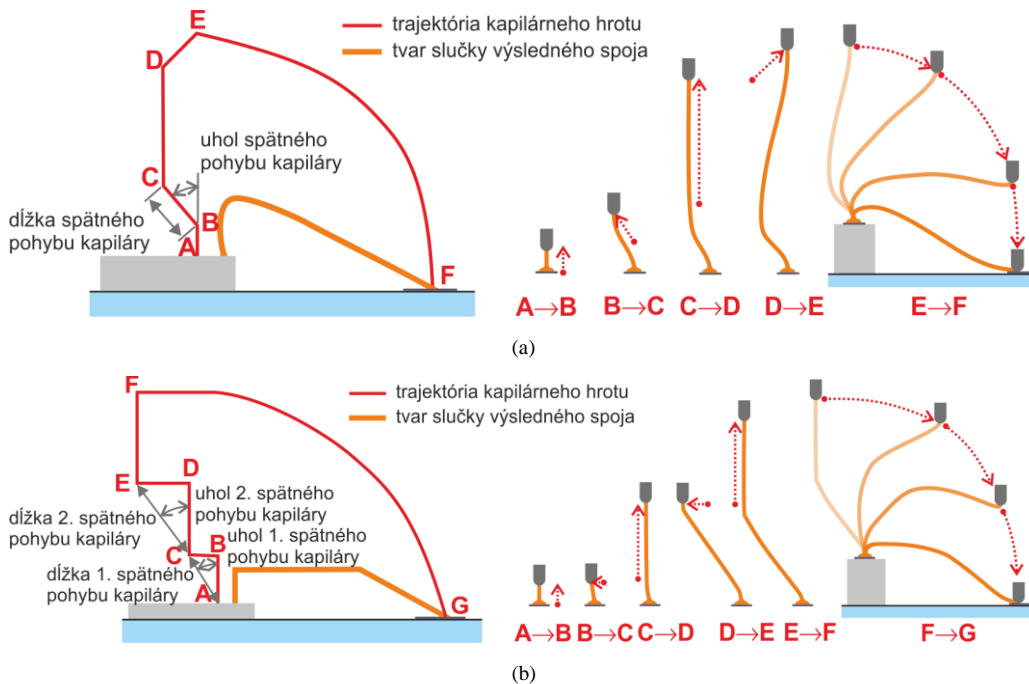
## II. VYTVÁRANIE KONTAKTOVANÝCH SPOJOV

Vo všeobecnosti rozoznávame z hľadiska fyzikálneho princípu, na akom sa kontaktovaný spoj vytvára, kontaktovanie termokompresné (tepelná energia + tlak), ultrazvukové (ultrazvuková energia) a termosonické (ultrazvuková energia + tepelná energia). Pri ultrazvukovom a termosonickom vytváraní kontaktov, ktoré tvoria v súčasnosti dominantné metódy kontaktovania, pôsobí na vytváraný kontakt aj tlak hrotu [2]. Z hľadiska vytvárania a tvarovania slučky kontaktovaného spoja je dôležité delenie kontaktovania podľa typu použitého kontaktovacieho hrotu. Rozoznávame kontaktovanie pomocou hrotu v tvare klínu (wedge bonding alebo aj wedge-wedge) a kontaktovanie pomocou kapiláry (ball bonding alebo aj ball-wedge). Kontaktovanie pomocou kapilárneho hrotu je rýchlejšia a robustnejšia metóda kontaktovania, napriek tomu v prípadoch, kde sa vyžadujú preporenia s nízkym profilom, sa využíva aj kontaktovanie pomocou hrotu v tvare klínu, ktoré navyše umožňuje aj vytvorenie neprerušeného stehu [5].



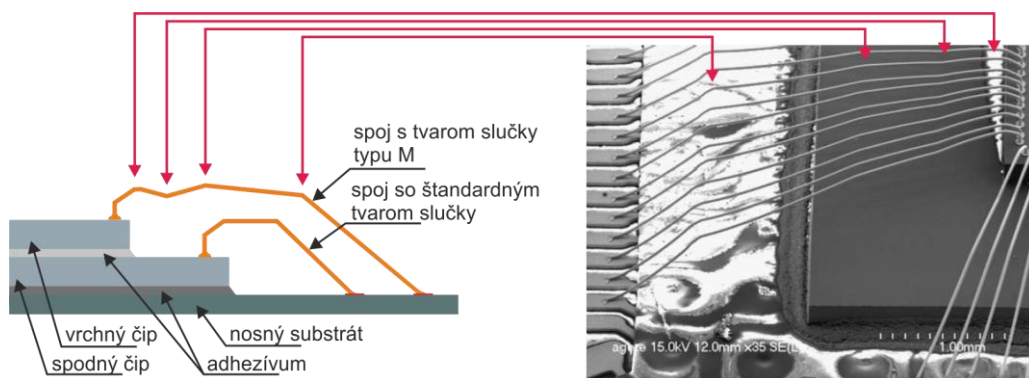
Obr. 1 Kapilárny kontaktovací hrot (a) a hrot s tvarom klínu (b)

Prvým dôležitým faktorom, ktorý vplyva na tvar výslednej slučky je trajektória, po ktorej sa pohybuje kontaktovací hrot. Problematike vplyvu trajektórie kapilárneho hrotu na tvar slučky je v súčasnosti venovaná väčšia pozornosť ako v prípade hrotu v tvare klinu, čo je spôsobené zrejme hlavne jeho častejším použitím v priemysle. Na Obr. 2 je znázornený vplyv trajektórie kapilárneho hrotu na výsledný tvar slučky kontaktovaného spoja s výslednou slučkou s trojuholníkovým tvarom (Obr. 2a) a s lichobežníkovým tvarom (Obr. 2b) [1][7].



Obr. 2 Vplyv trajektórie kapilárneho hrotu na výsledný tvar slučky kontaktovaného spoja: (a) slučka so štandardným trojuholníkovým tvarom (tiež tvaru T), (b) slučka s lichobežníkovým tvarom (podľa [1])

Moderné prepájacie a puzdriace technológie môžu mať na tvar slučky kontaktovaného spoja zvýšené požiadavky. Príkladom sú napr. technológie, pri ktorých sú polovodičové čipy ukladané na seba alebo technológia puzdrenia PiP (Package in Package), v ktorých je časté viacúrovňové kontaktovanie. V týchto prípadoch môžu byť okrem klasického lichobežníkového spoja používané aj zložitejšie tvary slučky kontaktovaného spoja (Obr. 3) [4]. Vo všeobecnosti je pre komplikovanejší tvar slučky výsledného spoja potrebná aj komplikovanejšia trajektória kontaktovacieho hrotu, ktorú je nutné definovať viacerými bodmi.



Obr. 3 Kontaktované spoje so slučkou tvaru M (podľa [4])

Vytváranie slučky v prípade hrotu tvaru klinu je komplikovanejšie, nakoľko na rozdiel od kapilárneho hrotu, pri ktorom kontaktovací drôtik prechádza driekom hrotu kolmo na kontaktovaciu plošku (Obr. 1a), pri hrote tvaru klinu je kontaktovací drôť vedený ku kontaktovej ploške pod rôznym sklonom (zväčša 30° a 60°). Menší sklon (30°), pod ktorým je kontaktovací drôť vedený ku kontaktovej ploške, je vhodnejší pre kontakty s vyššou slučkou, nakoľko poskytuje lepšiu ovládateľnosť umiestnenia pozície kontaktovaného spoja,

drôťik je ale v tomto prípade vystavený väčšiemu mechanickému namáhaniu. Naopak, pre kontakty s nižšou slučkou je vhodnejší 60° sklon, pri ktorom je mechanické namáhanie drôťika menšie [6][8].

Okrem trajektórie kontaktovacieho hrotu má pri vytváraní kontaktovaného spoja dôležitú úlohu materiál drôťika. Štandardne sa pri kontaktovaní používajú drôťiky na báze hliníka, zlata a najnovšie aj medi, zväčša s prídavkami iných prvkov pre upravenie ich mechanických vlastností. Ak chceme pochopiť vytváranie slučky kontaktovaného spoja, je potrebné charakterizovať zmeny materiálových vlastností vznikajúce pri rôznych deformačných rýchlostiach, najmä zmeny v Youngovom module, medze sklzu a ťažnosti. V prípade kontaktu vytváraného pomocou kapilárneho hrotu je potrebné vziať do úvahy aj dĺžku zóny, ktorá je vystavená tepelnému namáhaniu vplyvom EFO (electric flame-off) pri ktorom sa vytvára guľôčka pre prvý spoj [1][3].

### III. ZÁVISLOSŤ TVARU SLUČKY KONTAKTOVANÉHO SPOJA OD PARAMETROV KONTAKTOVANIA

Jednoduchý odhad tvaru slučky je možné realizovať pomocou vopred nameraných hodnôt tvaru slučky pri daných parametroch. Experimenty boli realizované na kontaktovacom zariadení Kulicke and Soffa model 4129 s hrotom tvaru klinu. Vedenie kontaktovacieho drôťika do klinu bolo pod uhlom 60°.

Odhladiac od nastavenia výšky prvého a druhého spoja, ktoré je závislé od umiestnenia kontaktovacích plôšok v konkrétnej aplikácii, umožňuje použité kontaktovacie zariadenie riadenie tvaru slučky pomocou 3 parametrov: loop, reverse a step back. Tvorba slučky kontaktovacieho spoja je realizovaná v niekoľkých krokoch (Obr. 4). Po vytvorení prvého kontaktu (na Obr. 4 v bode A) sa kontaktovací hrot posúva z bodu A do B, pričom výška posuvu hrotu  $H_{loop}$  je daná nastavením parametra „loop“. Nasleduje posun kontaktovacieho hrotu z bodu B do bodu C, ktoré je realizované buď automaticky pomocou parametra „step back“ (automatický posun je možný len do 0,5 mm) alebo manuálne (umožňuje vytvoriť dlhšie spoje). Pri vytváraní druhého kontaktu sa hrot najskôr posunie smerom nadol z bodu C do bodu D nad druhú kontaktovaciu plôšku a následne sa posunie smerom späť k prvému kontaktu (z bodu D do E). Veľkosť posunutia z bodu D do E ( $L_{reverse}$ ) umožňuje formovať slučku a je závislé na nastavení parametra „reverse“. Posledným krokom je vytvorenie druhého kontaktu (v bode F). Výsledná slučka má štandardný trojuholníkový tvar, ktorý môžeme zjednodušene charakterizovať jej dĺžkou ( $L$ ), výškou ( $H$ ) a vzdialenosťou vrcholu slučky od prvého kontaktu ( $L_1$ ).



Obr. 4 Vytváranie slučky kontaktovaného spoja na zariadení Kulicke and Soffa 4129

Na Obr. 5 je znázornený príklad závislosti výšky spoja od nastaveného parametra „loop“ a „reverse“, resp od  $H_{loop}$  a  $L_{reverse}$  pre kontaktovaný spoj s dĺžkou  $L = 0,9$  mm. Prvý aj druhý kontakt sa nachádzajú v jednej rovine. Podobné závislosti je možné vytvoriť aj pre iné dĺžky slučky  $L$  kontaktovaného spoja.

Vo všeobecnosti môžu byť namerané body výšky slučky spoja preložené a aproximované napr. rovinou so všeobecnou rovnicou:

$$H = a \cdot L_{reverse} - b \cdot H_{loop} - d$$

alebo kvadratickou plochou so všeobecnou rovnicou:

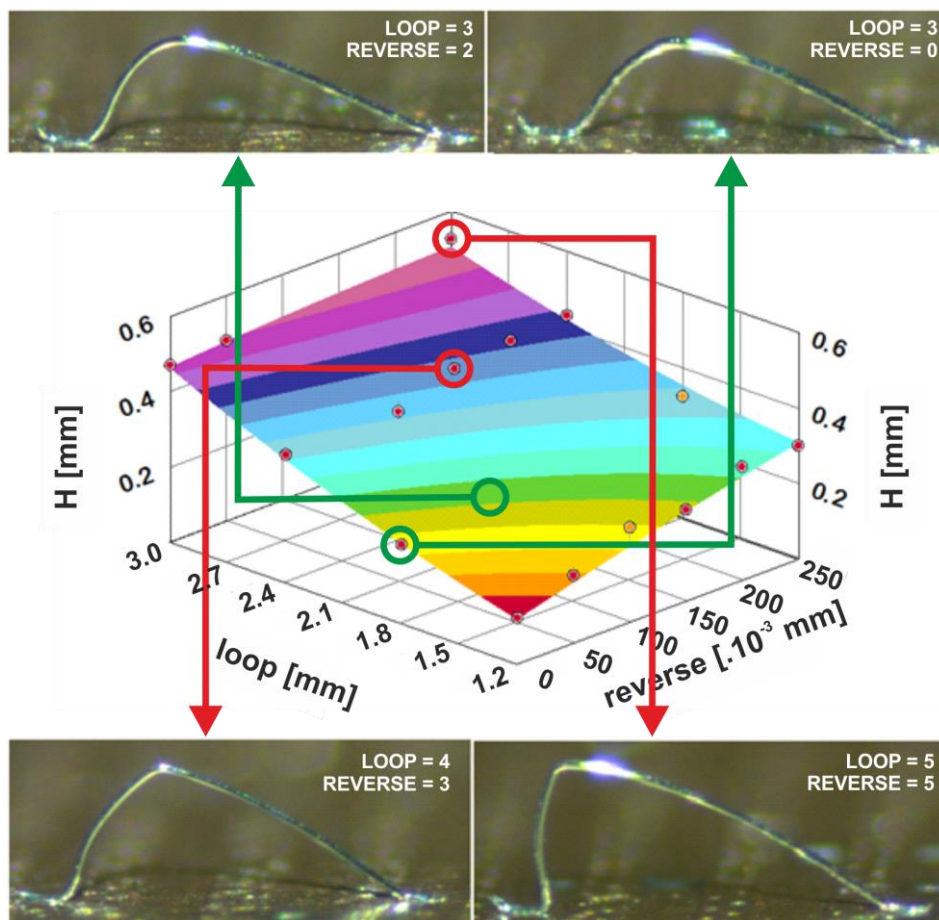
$$H = -a \cdot (L_{reverse})^2 - b \cdot (H_{loop})^2 - c \cdot L_{reverse} \cdot H_{loop} - d \cdot L_{reverse} - e \cdot H_{loop} - f$$

kde  $a, b, c, d, e$  a  $f$  sú parametre danej plochy, prípadne zložitejšou plochou. V grafe na Obr. 5 sú získané hodnoty preložené kvadratickou plochou, ktorá poskytuje oproti rovine presnejšie výsledky aj pri slučkách s inými dĺžkami  $L$ . Vypočítané koeficienty kvadratickej plochy pre slučky vytvorené v jednej rovine s rôznou dĺžkou  $L$  sú uvedené v Tab.1. Zahrnutím podobných

rovnic pre vzdialenosť vrcholu slučky kontaktovaného spoja od prvého kontaktu ( $L_1$ ) je možné získať sústavu dvoch rovníc z ktorých je možné následne vypočítať  $L_{reverse}$  a  $H_{loop}$ .

Tabuľka 1  
 Koeficienty kvadratickej plochy pre slučky vytvorené v jednej rovine s rôznou dĺžkou L

L (mm)	a (-)	b (-)	c (-)	d (-)	e (-)
0,3	0,010622	-0,019945	-0,091278	0,010311	0,037818
0,6	0,001573	-0,010697	-0,058526	0,010318	-0,039389
0,9	0,000950	-0,010779	-0,062661	0,009907	-0,039171
1,2	0,004902	-0,011329	-0,106089	0,009717	-0,005559
1,6	-0,002068	-0,010010	-0,088712	0,008993	0,004560
2	-0,004132	-0,007966	-0,126263	0,017608	-0,012391
2,4	-0,005789	-0,000426	-0,112347	0,010498	-0,036843
2,8	-0,043586	-0,003134	-0,154588	0,016724	-0,013598



Obr. 5 Závislosť výšky slučky kontaktovaného spoja (H) s dĺžkou  $L = 0,9$  mm od nastavených parametrov a mikroskopické snímky vybraných slučiek

Nevýhodou uvedenej metódy je samozrejme stále nový súbor vypočítaných parametrov plochy pre každú konkrétnu dĺžku H kontaktovaného spoja, čím sa získa súbor rovníc. Tento súbor by bol navyše platný len pre slučky v jednej rovine. Pre zahrnutie dĺžky H kontaktovaného spoja pre určenie parametrov, ako aj rozdielu vo výške prvého a druhého kontaktu (prvý kontakt slučky, je zväčša na strane čipu v osi z vyššie ako druhý kontakt vytvorený na kontaktovacej plôške), je preto potrebné realizovať viacrozmernú regresnú analýzu. Podobne je možné z nameraných dát získať závislosti pre polohu vrcholu od prvého spoja ( $L_1$ ).

#### IV. ZÁVER

Kontrola tvaru slučky kontaktovaného spoja je nevyhnutným predpokladom pre dosiahnutie optimálnej kvality kontaktovania, najmä pri stále sa zvyšujúcej integrácii a zvýšených

nárokoch pri puzdrení polovodičových čipov. Vysokofrekvenčné aplikácie sú zvlášť citlivé na tvar a ďalšie parametre kontaktovaného spoja, proces kontaktovania je v tejto oblasti kritickým krokom.

#### POĎAKOVANIE



*Tento článok bol vypracovaný s podporou projektu Vývoj unikátneho nízkoenergetického statického zdroja pre elektrosystémy, ITMS 26220220029, ktorý je spolufinancovaný zo štrukturálneho fondu EÚ ERDF v rámci výzvy OPV a V-2008/2.2/01-SORO a prioritnej osi 2 Podpora výskumu a vývoja.*

#### REFERENCES

- [1] D.S. Liu, Y.C. Chao, C.H. Wang, “Study of wire bonding looping formation in the electronic packaging process using the three-dimensional finite element method” in *Finite Elements in Analysis and Design*, Volume 40, Issue 3, January 2004, pp. 263–286, ISSN 0168-874X.
- [2] G. Harman, “*Wire Bonding in Microelectronics*”, 3rd Edition, New York: McGraw-Hill, 2010, ISBN: 978-0-07-164265-1.
- [3] Yu-Lung Lo, Chih-Chiang Tsao, “Wirebond profiles characterized by a modified linkage-spring model which includes a looping speed factor” in *Microelectronics Reliability*, Volume 42, Issue 2, February 2002, pp 285–291, ISSN 0026-2714.
- [4] Y.F. Yao, T.Y. Lin, K.H. Chua, “Improving the Deflection of Wire Bonds in Stacked Chip Scale Package (CSP)”, *Electronic Components and Technology Conference*, 2003. Proceedings. 53rd, May 27-30, 2003, IEEE 2013, pp. 1359–1363, ISSN 0569-5503.
- [5] D. J. Beck and A. C. Perez, “Wire Bond Technology - The Great Debate: Ball vs. Wedge”, In: *Advanced Packaging*, March, 2007.
- [6] Gaiser precision bonding tools, CoorsTek, Inc. <http://www.gaisertool.com/>
- [7] Yu-Lung Lo, Tei-Chen Chen, and Tieh-Lou Ho, “Design in Triangle-Profiles and T-Profiles of a Wirebond Using a Linkage-Spring Model”, *IEEE Transaction on Components and Packaging Technologies*, Vol. 24, No. 3, September 2001, pp. 457–467.
- [8] *Bonding Wedge Catalog* (Book style). Kulicke & Soffa, 2001.

# Analýza dynamických dejov v elektrizačnej sústave využitím SimPowerSystems

*Matúš KOLEJ, Dušan MEDVEĎ*

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
univerzita v Košiciach, Slovenská republika

matus.kolej@student.tuke.sk, dusan.medved@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá dynamickými dejmi v sústave vn, ktoré vznikajú pri skratoch, alebo následkom pripájania obnoviteľných zdrojov elektrickej energie. Analýza bola vykonaná použitím knižníc a nástrojov programu SimPowerSystems, toolboxom pre Simulink.

**Kľúčové slová** — SimPowerSystems, skrat, dynamický dej, simulácia, bioplynová elektrárňa, fotovoltaická elektrárňa

## I. ÚVOD

Výroba elektrickej energie sa v poslednom storočí orientovala hlavne na výstavbu a prevádzku veľkých a centralizovaných výrobní elektrickej energie ako sú napríklad jadrové, uhoľné a vodné elektrárne. Tieto elektrárne dodávajú elektrickú energiu hlavne do napäťovej hladiny zvn alebo vvn, pričom spotreba je umiestnená hlavne v sieťach vn a nn. Vidíme, že skladba energetických zdrojov sa každoročne mení v prospech obnoviteľných zdrojov elektrickej energie (OZE). Nehody, ktoré sa za posledné tri desaťročia odohrali v jadrových elektrárnach podnietili nevôľu laickej verejnosti ďalej podporovať jadrovú energetiku. Taktiež výstavba veľkých vodných diel s integrovanou výrobou elektrickej energie a tepelných elektrární sa stretáva s odporom ľudí hlavne vo vyspelých krajinách. Zvýšenie ekologického povedomia a obavy o životné prostredie medzi verejnosťou vedú politikov k presadzovaniu zákonov, ktoré podporujú ekologické zdroje elektrickej energie ako napríklad fotovoltaické elektrárne, bioplynové elektrárne, elektrárne na spaľovanie biomasy a iné, na úkor klasických zdrojov. Tieto zdroje, čo sa týka nižších výkonov, sú vo veľkej miere pripájané do napäťových hladín vn a nn. Výnimkou sú zdroje s vyššími inštalovanými výkonmi, ktoré sa pripájajú do vvn a veľké veterné farmy pripájané do zvn.

Tento príspevok sa zaoberá skúmaním dynamických javov v sústave vn s uvažovaním pripájania moderných obnoviteľných zdrojov elektrickej energie. Cieľom príspevku je pozorovať vplyv pripájania decentralizovaných OZE na zmenu skratových pomerov v uvažovanej sieti. Bude sa pozorovať ako ovplyvňujú OZE veľkosť skratových výkonov v mieste pripojenia.

## II. MODEL ELEKTRIZAČNEJ SÚSTAVY

Za účelom výskumu bol vytvorený detailný model siete vn v programe MATLAB. V nasledujúcej časti bude tento model stručne popísaný.

### A. Napájanie siete a transformátory

Napájanie systému zabezpečuje trojfázový zdroj „Three-Phase Source block“ na strane vvn, ktorý je v zapojení hviezda s priamo uzemneným neutrálnym bodom. Tento blok predstavuje napájanie uvažovanej 22 kV siete v elektrickej stanici z nadradenej 110 kV distribučnej siete.

Ďalším prvkom v schéme je transformátor. V simulácii sa nachádza 85 transformátorov typu „Three-Phase Transformer (Two Windings)“ rôznych výkonov a vnútorných zapojení vinutí. Sústavu 110 kV prepája so sústavou 22 kV jeden transformátor 110/22 kV s výkonom 40 MVA. Distribučných transformátorov 22/0,4 kV predstavuje 84 prvkov. (Pozn.: v praxi môžu byť vo vyhotovení ako stožiarové, kioskové, murované, kompaktné alebo vežové).



### B. Model vedenia

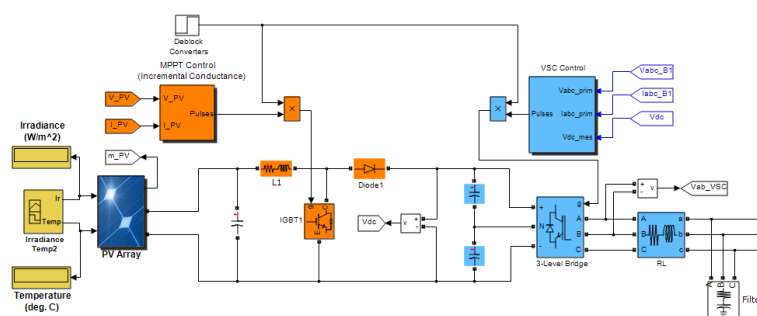
V simuláciách sa na modelovanie jednotlivých úsekov vedení používa blok „Three-Phase PI Section Line“, ktorý simuluje vedenie so sústrednými parametrami pomocou PI článku. Parametre vedenia  $R$ ,  $L$ ,  $C$  sa vypočítavajú zo súsledných a netočivých zložiek, ktoré berú do úvahy indukčné a kapacitné väzby medzi fázami navzájom a medzi fázami a zemou. Súsledné a netočivé zložky sú vypočítané pomocou špeciálneho nástroja na výpočet parametrov vzdušných vedení „Compute RLC Line Parameters“, ktorý ich vypočíta z parametrov vodiča a geometrického usporiadania vodičov.

### C. Model záťaže

Záťaže v simuláciách sú modelované pomocou bloku „Three-Phase Series RLC Load“. Jedná sa o model trojfázovej záťaže. Vstupnými hodnotami do bloku sú činný a jalový výkon.

### D. Model fotovoltaickej elektrárne

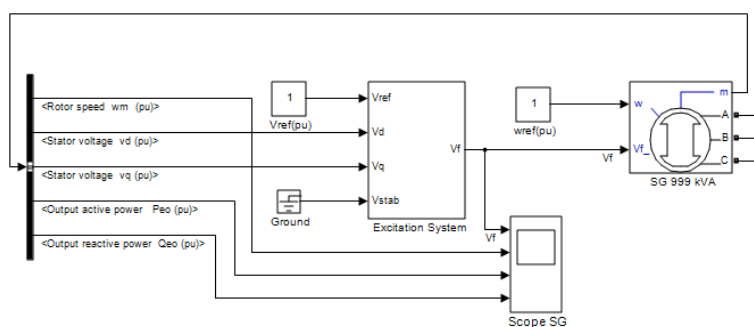
Na simuláciu fotovoltaickej elektrárne bol použitý model z prostredia SimPowerSystems, ktorý bol adekvátne upravený. Model fotovoltaickej elektrárne s výkonom 999 kW je vytvorený podľa technických podkladov k skutočnej fotovoltaickej elektrárne Gomboš 1 v okrese Košice – okolie.



Obr. 1 Model fotovoltaickej elektrárne

### E. Model bioplynovej elektrárne

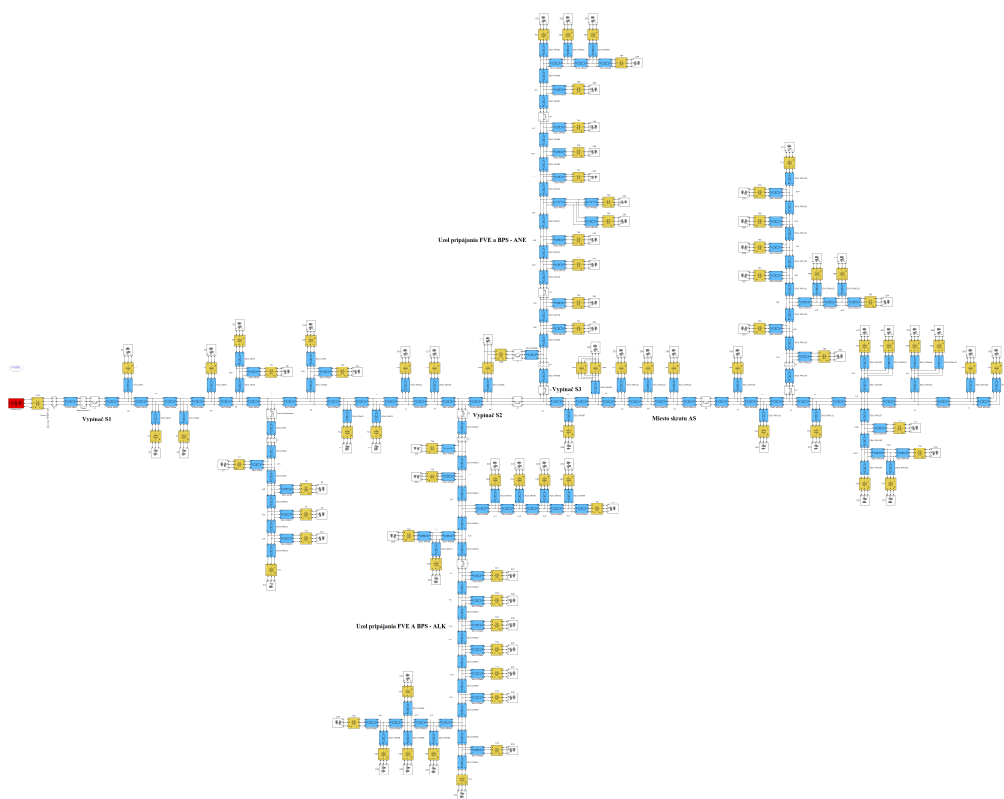
Model bioplynovej elektrárne je znázornený na Obr. 2. Synchronný generátor elektrárne je modelovaný pomocou bloku „Synchronous Machine“. Parametre modelov sú nastavené podľa katalógových údajov pre generátor AMG 0400CL04 (999 kVA). Budiaci systém je modelovaný pomocou bloku „Excitation System“, ktorý modeluje napäťový regulátor a budič.



Obr. 2 Model bioplynovej elektrárne

Model skúmanej časti siete môžeme vidieť na Obr. 3. Simulačný model je pomocou m-file súborov prepojený s databázou údajov, ktorá je uložená v špeciálnom súbore programu Excel. To poskytuje používateľovi nástroj na simulovanie rozličných problémov v ES. Prednosťou takéhoto riešenia je, že všetky údaje sú viditeľné a dostupné na jednom mieste. Dajú sa ľahko a prehľadne meniť, čo uľahčuje a zrýchľuje prácu.



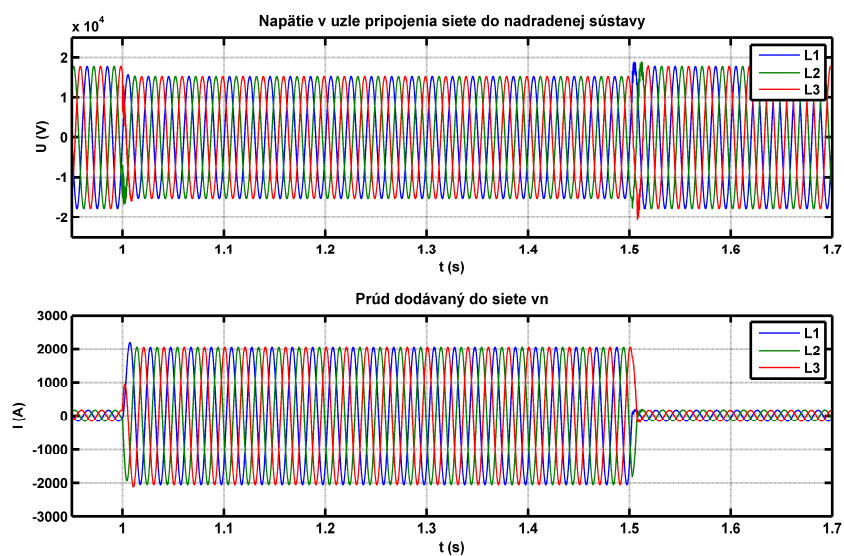


Obr. 3 Model siete v prostredí Simulink

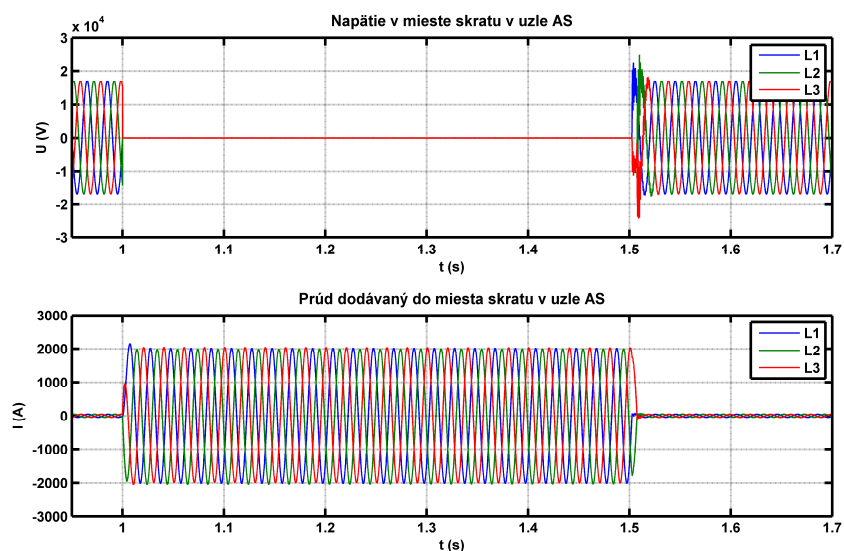
### III. VPLYV OZE NA SKRATOVÉ POMERY

Bol testovaný nasledovný scenár. V čase 1 s nastal na vedení v uzle AS trojfázový kovový skrat, ktorý zanikol v čase 1,5 s. Do uzla ANE boli postupne pripojené OZE: fotovoltaická elektrárňa FVE 999 kW a bioplynová elektrárňa BPS 999 kVA. Výsledné priebehy (Obr. 5 až Obr. 9) a zistené hodnoty (Tabuľky 1 a 2) sú porovnané s referenčnými priebehmi a hodnotami zistenými pre sieť bez pripojených OZE (Obr. 4).

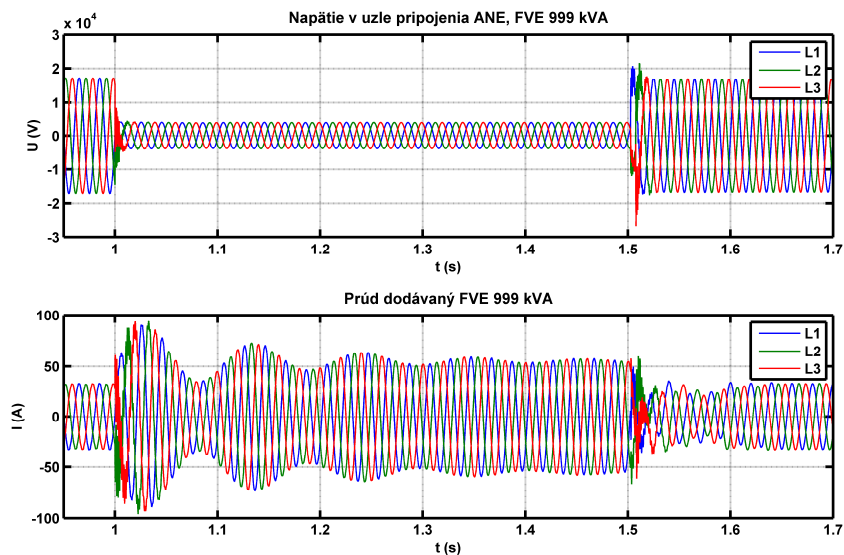
Cieľom simulácií bolo overiť vplyv pripojených OZE na skratové pomery a zmenu skratových výkonov v sieti bez OZE a s OZE.



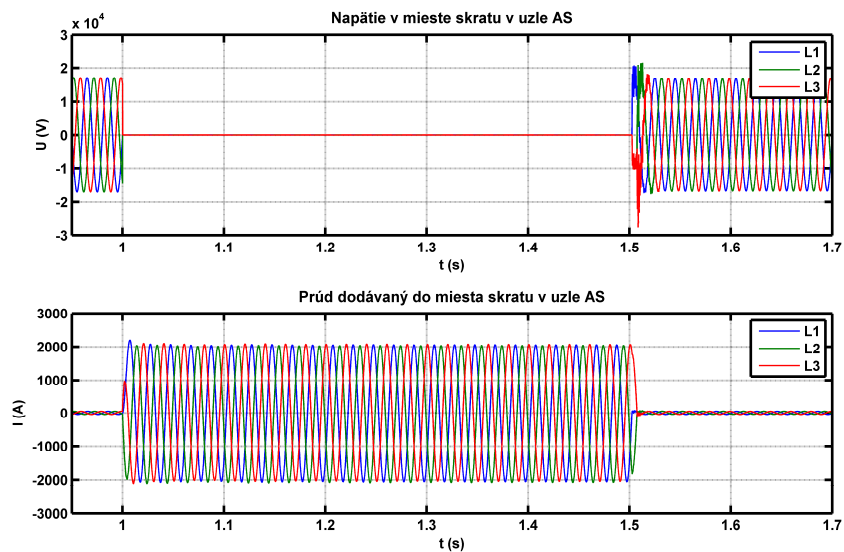
Obr. 4 Napätia a prúdy v mieste pripojenia siete do nadradenej sústavy bez pripojených OZE



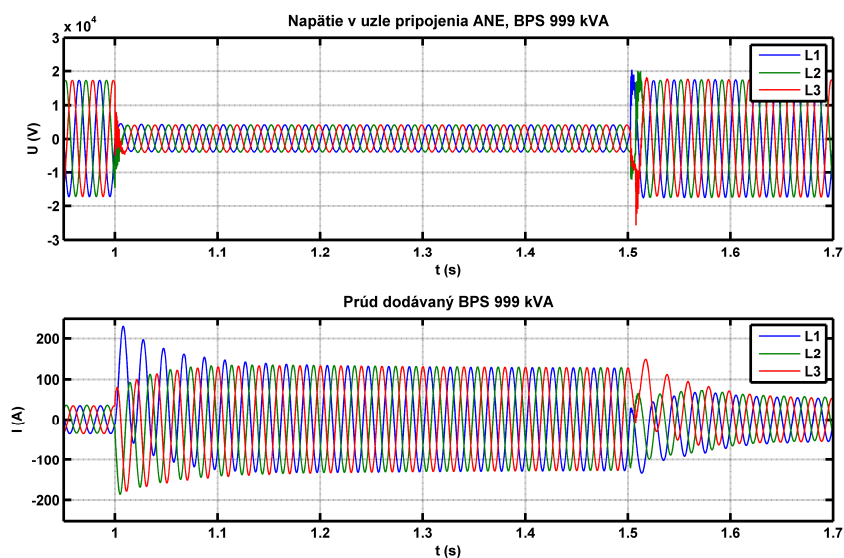
Obr. 5 Napätia a prúdy v mieste skratu (uzol AS) bez pripojených OZE



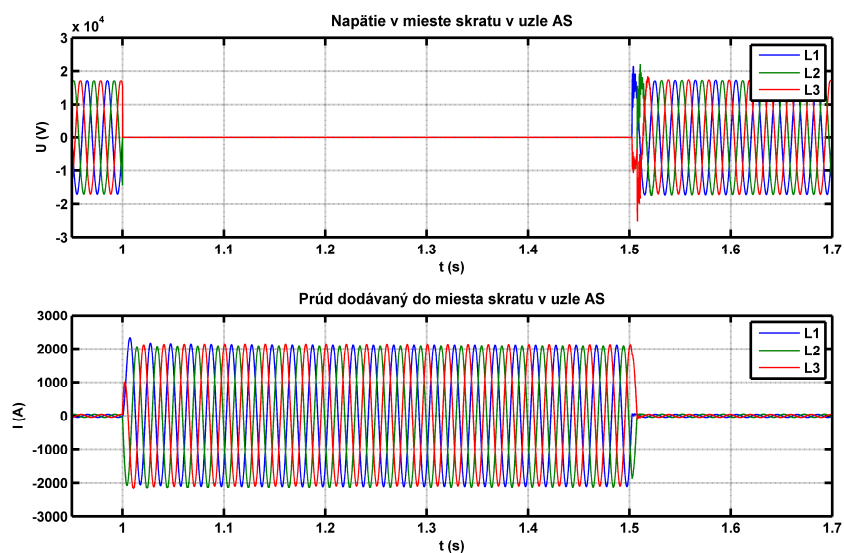
Obr. 6 Napätia a prúdy v mieste pripojenia FVE 999 kVA v uzle ANE



Obr. 7 Napätia a prúdy v mieste skratu (uzol AS) v sieti s pripojenou FVE 999 kVA



Obr. 8 Napätia a prúdy v mieste pripojenia BPS 999 kVA v uzle ANE



Obr. 9 Napätia a prúdy v mieste skratu (uzol AS) v sieti s pripojenou BPS 999 kVA

Tabuľka 1  
 Hodnoty fázových napätí (amplitúdy) v mieste skratu a v mieste pripojenia OZE (v uzle ANE) pri trojfázovom kovovom skrate

Prípado		Sieť bez OZE	Sieť s FVE 999 kVA	Sieť s BPS 999 kVA
Veličina v mieste pripojenia siete alebo OZE (uzol ANE)				
$U_{\text{predskratom}}$	[V]	17843	17234	17335
$U_{\text{počas skratu}}$	[V]	15393	3865	4082,5
$U_{\text{ust.poskrate}}$	[V]	17846	16868	17331
v mieste skratu (uzol AS)				
$U_{\text{predskratom}}$	[V]	16984	17113	17218
$U_{\text{počas skratu}}$	[V]	2,02	2,06	2,12
$U_{\text{ust.poskrate}}$	[V]	16983	16822	17215

Tabuľka 2  
 Hodnoty fázových prúdov (amplitúdy) v mieste skratu a v mieste pripojenia OZE (v uzle ANE) pri trojfázovom kovovom skrate

Prípado		Sieť bez OZE	Sieť s FVE 999 kVA	Sieť s BPS 999 kVA
Veličina v mieste pripojenia siete alebo OZE (uzol ANE)				
$I_{predskratom}$	[A]	165,8	33,08	34,46
$i_p$	[A]	2193	96	229,28
$I_{k3,max}''$	[A]	2055	56,5	126,65
$I_{k3}''$	[A]	1453	39,95	89,56
$I_{ust,postrate}$	[A]	165,5	32,95	33,4
v mieste skratu (uzol AS)				
$I_{predskratom}$	[A]	48,5	49	45
$i_p$	[A]	2155	2201	2332
$I_{k3,max}''$	[A]	2020	2062	2115
$I_{k3}''$	[A]	1428	1458	1495,5
$I_{ust,postrate}$	[A]	48,3	47,9	44,7

Porovnaním priebehov a hodnôt v tabuľkách zisťujeme, že v prípade pripojenia BPS dôjde k väčšiemu zvýšeniu hodnoty skratového prúdu ako pri FVE. Je to spôsobené tým, že menič FVE je obmedzený maximálnym dovoleným prúdom, ktorý môže dodať. BPS je teda väčším zdrojom skratového prúdu a skratového výkonu. Hodnoty zistených napätí a prúdov sú v Tabuľke 1 a 2.

#### IV. ZÁVER

Príspevok rieši vplyv OZE na skratové pomery v sieti vn. Môžeme skonštatovať, že synchronne generátory v BPS prispievajú do skratu väčším poruchovým prúdom ako FVE. Nižšie príspevky FVE sú dané obmedzujúcim dovoleným prúdom meničov FVE. S vyšším skratovým prúdom v mieste pripojenia súvisí aj zvýšenie skratového výkonu. Skratový výkon určuje tvrdosť siete a môžeme ho považovať za parameter charakterizujúci kvalitu elektrickej energie v danom mieste. Pripájanie BPS do siete vn zvyšuje skratový výkon v mieste pripojenia a teda môžeme tvrdiť, že pripájanie BPS má pozitívny dopad na kvalitu elektrickej energie. Príkladom môže byť flicker, ktorý je závislý od skratového výkonu a s rastúcim skratovým výkonom klesá. Vyšší skratový výkon tiež obmedzuje vplyv vyšších harmonických.

#### POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied grantom VEGA č. 1/0388/13.

#### REFERENCIE

- [1] KOLEJ, Matúš: *Analýza dynamických dejov v elektrizačnej sústave využitím SimPowerSystems*. Diplomová práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2014.
- [2] *Help programu Simulink*.
- [3] *Help programu SimPowerSystem*.
- [4] *Fotovoltaická elektrárň Gomboš*. [online]. [cit. 2013-12-1]. Dostupné na internete: <<http://www.enviroportal.sk/sk/eia/detail/fotovoltaicka-elektararn-gom-bos-1>>.
- [5] GIROUX, Pierre a kol.: *Grid-Connected PV Array*. [online]. [cit. 2013-15-1]. Dostupné na internete: <<http://www.mathworks.com/matlabcentral/fileexchange/34752-grid-connected-pv-array>>.
- [6] ABB: *Low Voltage Generators for Industrial Applications*. [online]. [cit. 2014-04-18]. Dostupné na internete: <<http://www.abb.sk/product/seitp322/6c067ef0005eb928c12576160033c3d8.aspx>>.

# Lakovací robot

<sup>1</sup>Peter GIROVSKÝ, <sup>2</sup>Jaroslava Žilková, <sup>3</sup>Marek VACEK, <sup>4</sup>Tomáš GNEVEK

<sup>1,2,3</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

<sup>1</sup>peter.girovsky@tuke.sk, <sup>2</sup>jaroslava.zilkova@tuke.sk, <sup>3</sup>marek.vacek@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá návrhom riadenia a simuláciou lakovacieho robota. V rámci článku je popísaná priama a inverzná kinematická úloha lakovacieho robota. Výsledok priamej a inverznej kinematickej úlohy bol použitý pre návrh riadenia virtuálneho lakovacieho robota. Pre návrh a simuláciu virtuálneho lakovacieho robota bol použitý program Matlab.

**Kľúčové slová** — robot, lakovanie, virtuálny model, simulácia

## I. ÚVOD

Témy robotických systémov a virtuálnej reality sú v dnešnej dobe veľmi rozšírené. Priemyselné roboty, ako aj simulácie vo virtuálnej realite nachádzajú čoraz častejšie uplatnenie pri čoraz rôznorodejších aplikáciách. V tomto príspevku chceme prezentovať spojenie týchto dvoch vysoko aktuálnych tém v jednej práci simulácie lakovacieho robota vo virtuálnej realite.

V súčasnej dobe predstavujú robotické systémy rozvinuté autonómne systémy ktoré výkonne a efektívne spolupracujú s výrobnými systémami v priemysle, ale postupne nachádzajú aj široké uplatnenie ostatných nevýrobných resp. nepriemyselných odvetviach. Ich čoraz častejšie vyžívanie je dôsledkom snahy zvýšiť produktivitu výrobného procesu, zníženie nákladov a zvýšenie presnosti výroby.

Taktiež virtuálna realita sa čoraz častejšie využíva najmä v technických odvetviach ako nástroj pre riešenie mnohých špecifických problémov. Samotný pojem virtuálna realita možno chápať ako nejaké prostredie vytvorené pomocou počítačových prostriedkov, ktoré umožňuje prácu v trojrozmernom priestore. Samotná tvorba virtuálnej reality v sebe zahŕňa postupy v oblasti počítačovej grafiky, kde ide predovšetkým o tvorbu trojrozmerných modelov, manipuláciu s týmito modelmi, pohyb v trojrozmernom priestore a jeho súčasné zobrazenie v reálnom čase. Najväčší dôraz, kladený na simuláciu vo virtuálnom prostredí, je okamžitá odozva ktorejkoľvek činnosti vo virtuálnom prostredí. Aby bola virtuálna realita čo najviac podobná skutočnému svetu, musia byť splnené nasledovné vlastnosti:

- všetky deje sa musia vykonávať v reálnom čase,
- virtuálny priestor spolu s objektmi musia byť trojrozmerné, alebo aspoň vytvárať takúto ilúziu,
- užívateľ musí vedieť vstupovať do virtuálneho sveta, ovplyvňovať ho a manipulovať s jeho jednotlivými časťami.

## II. KINEMATIKA LAKOVACIEHO ROBOTA

Riešenie kinematiky ramena robota predstavuje riešenie priamej a inverznej kinematickej úlohy a je nevyhnutné pre generovanie riadiacich trajektórií pohybu ramena robota.

### A. Priama kinematická úloha

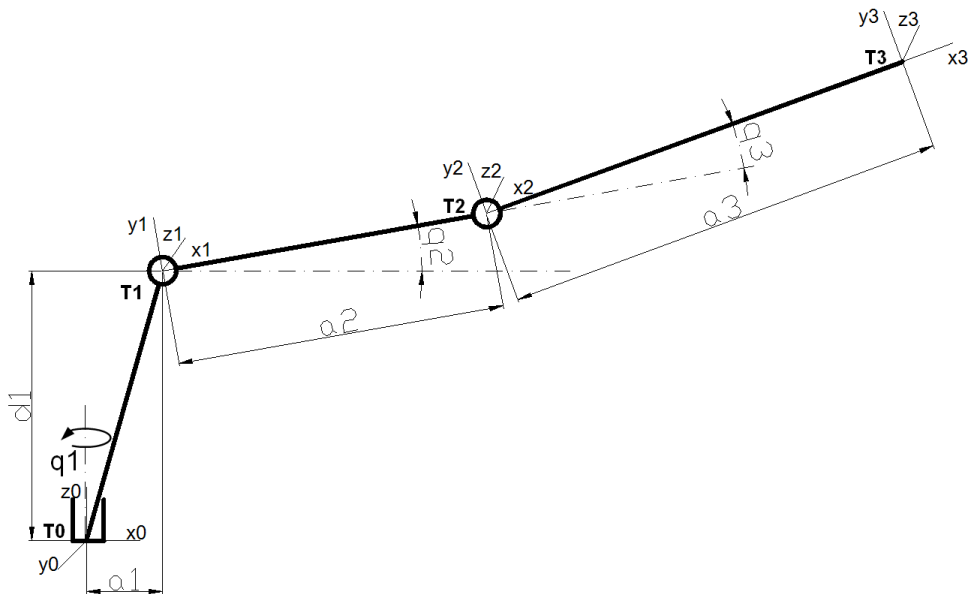
Najčastejšie používaným spôsobom voľby súradnicových systémov pre riešenie kinematickej úlohy je spôsob podľa Denavit-Hartenberga. Výsledná transformačná matica priamej kinematickej úlohy, vyjadruje závislosť medzi priestorovými súradnicami koncového bodu ramena robota a referenčným súradnicovým systémom robota.

Vzťahy medzi jednotlivými súradnicovými systémami môžeme popísať na základe transformačných matic podľa kinematickej štruktúry znázornenej na obr. 1. Získavame tak tri transformačné matice, ktoré pozostávajú z matic charakterizujúcich pootočenie a posuv v jednotlivých osiach súradnicových systémov. Vynásobením týchto troch transformačných matic dostávame výslednú transformačnú maticu (1).

$${}_{0,0} = {}_{1,0} \cdot T_{2,1} \cdot T_{,2} = \begin{bmatrix} l & m & n & r \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} l_x & m_x & n_x & r_x \\ l_y & m_y & n_y & r_y \\ l_z & m_z & n_z & r_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Kde vektory  $l$ ,  $m$ ,  $n$  určujú orientáciu a vektor  $r$  určuje polohu koncového bodu ramena robota. Jednotlivé zložky týchto vektorov sú:

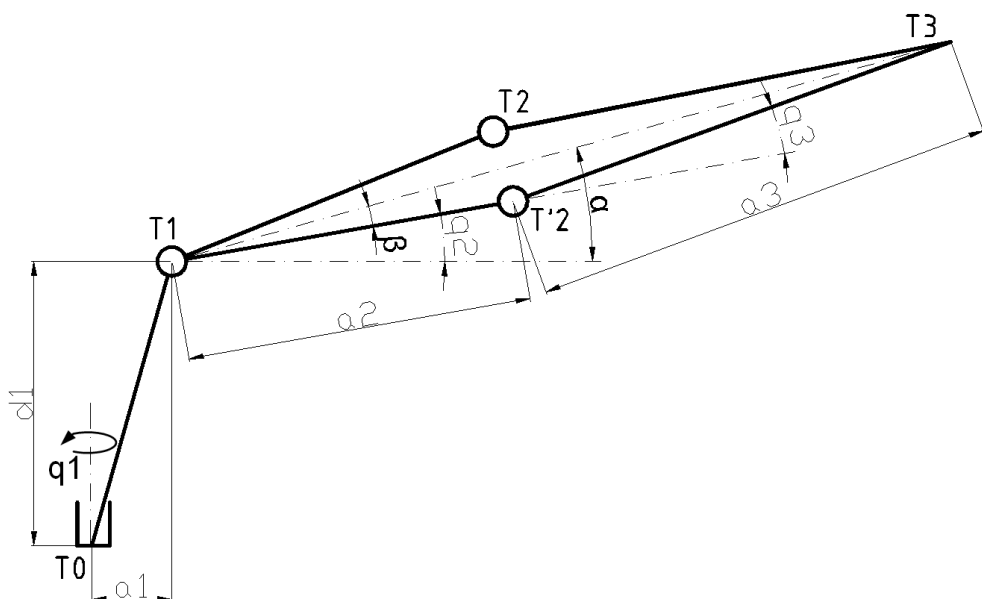
$$\begin{aligned} l_x &= \cos(q_2 + q_3) \cdot \cos(q_1) & m_x &= -\sin(q_2 + q_3) \cdot \cos(q_1) & n_x &= \sin(q_1) \\ l_y &= \cos(q_2 + q_3) \cdot \sin(q_1) & m_y &= -\sin(q_2 + q_3) \cdot \sin(q_1) & n_y &= -\cos(q_1) \\ l_z &= \sin(q_2 + q_3) & m_z &= \cos(q_2 + q_3) & n_z &= 0 \end{aligned}$$



Obr. 1 Kinematická schéma lakovacieho robota.

### B. Inverzná kinematická úloha

Riešenie inverznej úlohy kinematiky robota, ktorého kinematická štruktúra je znázornená na Obr. 1 vychádza z riešenia rovinného manipulátora s dvoma stupňami voľnosti, ktorý vznikne po rozdelení úlohy na rotáciu roviny okolo osi  $z_0$  a na rovinný pohyb ramena v tejto rovine. Schéma upravenej kinematickej štruktúry je zobrazená na obr. 2.



Obr. 2 Schéma upravenej kinematickej štruktúry pre výpočet inverznej kinematiky.

Riešením inverznej kinematickej úlohy získame rovnice inverznej kinematiky robota. Po dosadení hodnôt súradníc  $x$ ,  $y$ ,  $z$  a rozmerov ramien robota dostaneme hodnoty natočení jednotlivých kĺbov.

Hodnotu  $q_1$  vypočítame zo súradníc koncového bodu  $x$  a  $z$ , ktoré tvoria odvesny pravouhlého trojuholníka na horizontálnej rovine.

$$q_1 = \arctan\left(\frac{z}{x}\right) \quad (2)$$

Hodnoty  $q_2$  a  $q_3$  vypočítame podľa upravenej kinematickej schémy zobrazenej na Obr. 2, kde uvažujeme len časť ramena od člena  $T_1$  po  $T_3$ .

$$q_2 = \arctan\left(\frac{y}{x}\right) + \arccos\left(\frac{a_2^2 - a_3^2 + c^2}{2 \cdot c \cdot a_2}\right) \quad (3)$$

$$q_3 = \pm \arccos\left(\frac{x^2 + y^2 - (a_2^2 + a_3^2)}{2 \cdot a_2 \cdot a_3}\right) \quad (4)$$

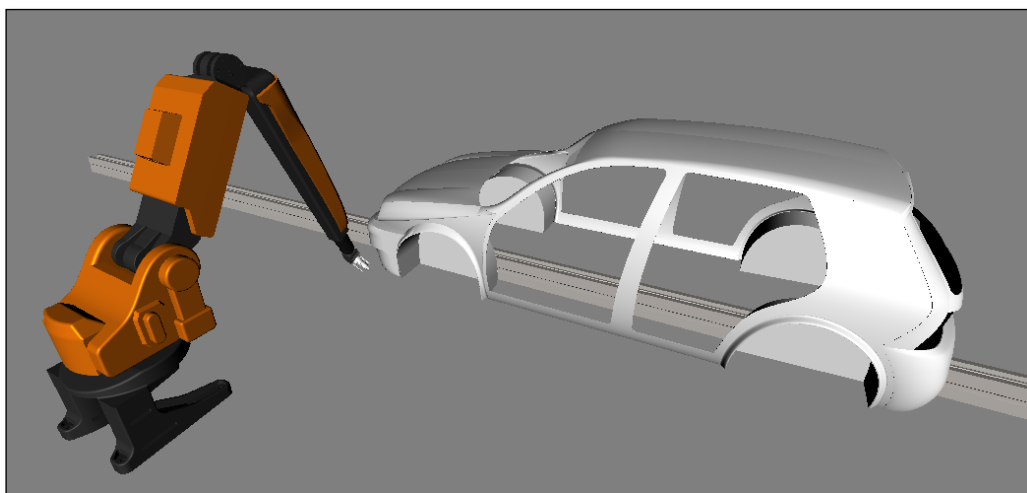
### III. SIMULÁCIA LAKOVACIEHO ROBOTA

#### A. Virtuálny model lakovacieho robota

Pre simuláciu lakovacieho robota bolo najskôr potrebné zostrojiť samotný virtuálny model lakovacieho robota a pracovnej stanice. Model lakovacej celej stanice pozostáva z troch samostatných celkov:

- 3D model lakovacieho robota,
- časť dopravníkového pásu lakovacej stanice,
- karoséria automobilu.

Na obr. 3 je zobrazený výsledný model lakovacej stanice.



Obr. 3 3D model lakovacej stanice.

#### B. Riadenie simulácie lakovacej stanice

Navrhnutá lakovacia stanica sa riadi pomocou dvoch matic, z ktorých jedna riadi pohyby robota a druhá ovláda pozíciu karoséria automobilu. Matica ovládania robota je zložená z ôsmich stĺpcov, ktoré predstavujú:

1. vektor času, v ktorom nastávajú zmeny jednotlivých pohybov,
2. súradnica  $x$  pozície koncového bodu robota,
3. súradnica  $y$  pozície koncového bodu robota,
4. súradnica  $z$  pozície koncového bodu robota,
5. natočenie prvého kĺbu zápästia trysky,
6. natočenie druhého kĺbu zápästia trysky,
7. natočenie tretieho kĺbu zápästia trysky,
8. natočenie samotnej trysky.

Keďže do stĺpca 2-4 sa zadávajú hodnoty súradníc pozície koncového bodu robota, je potrebné pomocou inverznej kinematiky tieto hodnoty prepočítať na hodnoty natočenia prvého, druhého a tretieho kĺbu. Do stĺpcov 5-8 sa zadávajú uhly natočenia kĺbov zápastia a trysky v stupňoch. Výsledná matica riadenia pozostáva z vektora času a siedmych vektorov natočenia kĺbov v radiánoch.

Matica ovládania pozícia karosérie automobilu je zložená z piatich stĺpcov a je ňou možné ovládať pozíciu karosérie v smere osi z, pozíciu karosérie v smere osi y (výška), natočenie karosérie okolo osi y a natočenie karosérie okolo osi z. Jednotlivé stĺpce matice riadenia karosérie predstavujú:

1. vektor času, v ktorom nastávajú zmeny jednotlivých pohybov,
2. pozícia v smere osi z (posuv vpred/vzad),
3. pozícia v smere osi y (výška),
4. natočenie okolo osi z,
5. natočenie okolo osi y.

#### IV. ZÁVER

V tomto príspevku sme sa venovali vytvoreniu 3D modelu lakovacej stanice automobilov v programe MATLAB, riešeniu priamej a inverznej kinematiky lakovacieho robota a jeho riadeniu pomocou MATLABu.

Výsledkom tejto práce bolo vytvorenie animácie lakovacej stanice. V animácii prebieha simulácia lakovania pomocou lakovacieho robota, ktorý postupne nalakuje celú karosériu automobilu z vonkajšej strany. Karoséria počas simulácie nemení svoju farbu z dôvodu obmedzených možností animácie v programe MATLAB. Pre zlepšenie tejto simulácie, by sme navrhovali zaradiť do lakovacej stanice druhého lakovacieho robota na pravú stranu karosérie automobilu, ktorý by pracoval súčasne s prvým, čo by viedlo k zrýchleniu procesu lakovania a taktiež by odpadla potreba otáčania karosérie pre nalakovanie jej druhej strany.

#### POĎAKOVANIE

Táto práca bola podporená projektom KEGA - 042TUKE-4/2012.

„Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: **Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií**, kód ITMS: **26220220182**, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.“



„Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ“

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.

#### LITERATÚRA

- [1] R. Grepl, Kinematika a dynamika mechatronických systémů. Brno: CERM, 2007. 158 s. ISBN 978-80-214-3530-8
- [2] I. Havel, Robotika: úvod do teorie kognitivních robotů. Praha: SNTL, 1980.
- [3] F. Šolc, L. Žalud, Robotika, Brno, 2002.
- [4] A Palko, J. Smrček, Robotika: koncové efekторы pre priemyselné a servisné roboty. Košice: TU, 2004. 272 s. ISBN 80-8073-218-3
- [5] J. Žára, VRML 97 – Laskavý průvodce virtuálními světy. Brno: Computer Press, 1999, 238 s., ISBN 80-7226-143-6



# Elektronické riadenie stability vozidla

<sup>1</sup>Jaroslava ŽILKOVÁ, <sup>2</sup>Peter GIROVSKÝ, <sup>3</sup>Jaroslav KUCHAR, <sup>4</sup>Marek VACEK

<sup>1,2,4</sup> Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>jaroslava.zilkova@tuke.sk, <sup>2</sup>peter.girovsky@tuke.sk, <sup>4</sup>marek.vacek@tuke.sk

**Abstract** — Príspevok sa zaoberá návrhom riadenia ESP systému vozidla. Na základe modelu vozidla sú v programe CarSim testované a porovnané tri modely vozidla, bez ESP systému, s klasickým ESP systémom a riešenie s fuzzy ESP systémom. Výsledky testovania týchto troch modelov vozidiel potvrdzujú, že elektronický regulátor stability podstatne zvyšuje bezpečnosť prevádzky vozidla v cestnej premávke, pričom vozidlo s navrhnutým fuzzy riadením ESP dosiahlo lepšie výsledky než vozidlo s klasickým ESP systémom.

**Keywords** — elektronické riadenie stability, fuzzy riadenie, model vozidla,

## I. ÚVOD

Rozširovanie cestnej dopravy má stále stúpajúcu tendenciu a okrem nesporných výhod prináša aj čoraz náročnejšie požiadavky na dopravu a jej bezpečnosť. Bezpečnosť dopravy sa v súčasnosti stala nielen vážnym dopravným, spoločenským, ale aj ekonomickým problémom. Dopravná nehodovosť sa spája s veľkými materiálnymi škodami, trvalými následkami na zdraví obyvateľov a aj so stratami na ľudských životoch.

Pri vývoji nových automobilov sa kladie veľký dôraz na splnenie bezpečnostných podmienok, ktoré sú predpísané platnou legislatívou jednotlivých krajín, a keďže bezpečnostné prvky v automobile zvyšujú aj jeho konkurencioschopnosť, sú vyvíjané stále nové prvky zvyšujúce bezpečnosť posádky automobilu a minimalizujúce následky nehôd. Bezpečnostné prvky rozdeľujeme do dvoch skupín na aktívne, ktoré svojim pôsobením znižujú pravdepodobnosť vzniku nehody, a pasívne, ktoré pri vzniku nehody minimalizujú jej následky. Vzhľadom na stále rastúce požiadavky zvyšovania bezpečnosti automobilov sa neustále zintenzívňuje výskum nových aktívnych aj pasívnych bezpečnostných prvkov a ich inovácií.

K pasívnym bezpečnostným prvkom okrem robustnej karosérie, patria bezpečnostné pásy a zostava airbagov, ktorá chráni cestujúcich zo všetkých strán, aktívne opierky hlavy, bočné výstupy, detské poistky, FPS (Fire Protection System Safety) systém zabraňujúci vznieteniu vozidla, parkovací asistent a ďalšie.

Najdôležitejšími prvkami aktívnej bezpečnosti sú účinné brzdy umožňujúce bezpečné spomalenie alebo zastavenie vozidla, dobrý výhľad z vozidla, dobré pneumatiky, presné a spoľahlivé riadenie, kvalitné tlmiče zabezpečujúce dostatočný kontakt pneumatík s vozovkou a osvetlenie vozidla. Medzi ďalšie prvky aktívnej bezpečnosti patria moderné elektronické systémy ako napr.: ABS, ESP, TCS, EBA, ACC a ďalšie.

Predložený článok sa zaoberá návrhom a verifikáciou fuzzy riadenia elektronického systému stability vozidla pomocou programu CarSim a Matlab/Simulink.

## II. ELEKTRONICKÝ STABILIZAČNÝ PROGRAM ESP (ELECTRONIC STABILITY PROGRAM)

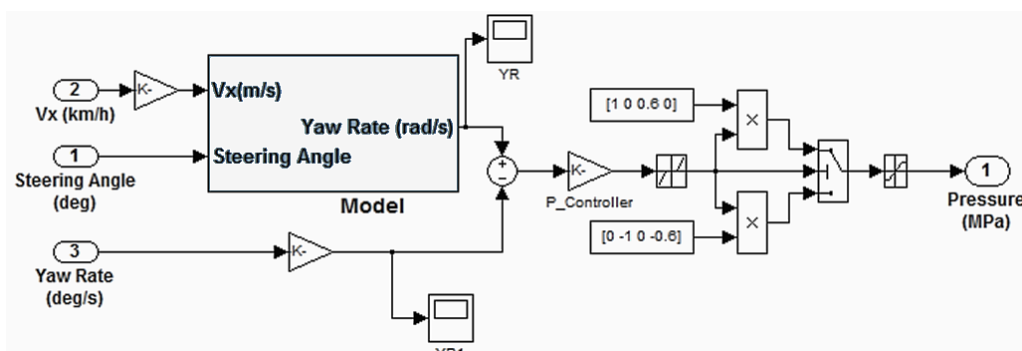
Zabezpečenie stability vozidla, ktorým sa v príspevku zaoberáme, patrí k aktívnym bezpečnostným prvkom, ktoré hlavne v kritických situáciách môže zabrániť dopravnej nehode.

Elektronický stabilizačný program neustále monitoruje situáciu na vozovke a zabezpečuje stabilitu vozidla pri vyhábacích manévroch, pri prejazdoch zatáčkami a v rôznych hlavne kritických situáciách. Ak ESP zaznamená odchýlku od zvolenej jazdnej stopy, automaticky upraví smer auta na vozovke. ESP je rozšírením protiblokovacieho systému bŕzd ABS a protipreklzového systému ASR, ktoré ovládajú sklz a preklz pneumatiky len v pozdĺžnom smere vozidla. Elektronický stabilizačný systém ESP reguluje sklz pneumatiky aj v priečnom

smere vozidla, znižuje nebezpečenstvo šmyku pri brzdení, zrýchlení aj pri voľnej jazde a pri prejazde zákrutou zvyšuje stabilitu držania stopy vozidla. Systém zároveň sleduje aj reakciu vodiča a mikroprocesor behom niekoľkých milisekúnd vyhodnotí situáciu a určí, ktoré koleso a akou silou je potrebné pribrzdiť a ako je potrebné znížiť hnací moment motora, aby bolo vozidlo opäť stabilizované.

#### A. Elektronické riadenie stability

Pri pretáčavom pohybe, ak hrozí vybočenie zadnej časti vozidla, sa pribrzďujú kolesá na vonkajšej strane zákruty, pričom na predné vonkajšie koleso pôsobí najväčšia brzdná sila. Pri nedotáčavom pohybe pôsobí najväčšia brzdná sila na zadnom vnútornom kolese a korekcia sa prevedie pribrzdením kolies na vnútornej strane zákruty. Systém zároveň sleduje aj reakciu vodiča a mikroprocesor počas niekoľkých milisekúnd vyhodnotí situáciu a určí, ktoré koleso a akou silou je potrebné pribrzdiť a ako je potrebné znížiť hnací moment motora, aby bolo vozidlo opäť stabilizované. Princíp elektronického regulátora stability (obr.1), ktorý je prístupný programe CarSim spočíva v tom, že v modeli vozidla, sa pomocou rýchlosti vozidla a uhla natočenia kolies vypočíta žiadaná miera stáčania, ktorá je porovnaná so skutočnou mierou stáčania zo snímača na vozidle. Rozdiel medzi žiadanou a skutočnou mierou stáčania tvorí regulačnú odchýlku, ktorá vstupuje do regulátora (obr.1) za ktorým nasleduje funkčný blok dead pásma s rozsahom od -2 do 2. Ak je výstup regulátora v tomto rozsahu, potom výstup bude nulový. Ak hodnota nie je v tomto rozsahu, v závislosti či je výstupná hodnota kladná alebo záporná, logika riadenia bŕzd určí, ktorá strana vozidla bude pribrzďovaná. Ak je hodnota výstupu regulátora kladná, bude násobená blokom konštánt [1 0 0.6 0] čo znamená, že brzdí pravé predné (1) a pravé zadné (0.6) koleso a ak je hodnota výstupu regulátora záporná, bude násobená blokom konštánt [0 -1 0 -0.6], čo znamená, že brzdia kolesá na ľavej strane.



Obr. 1 Schéma regulátora stability.

Dôvodom rozdielnej hodnoty medzi zadným a predným kolesom je zmena ťažiska pri brzdení, keď sa ťažisko vozidla prenáša na predné kolesá a zadné kolesá sú odľahčené, takže predné brzdy musia brzdiť viac ako zadné. Výstupom regulátora stability je hodnota určujúca potrebný tlak valca na brzdový kotúč jednotlivých kolies.

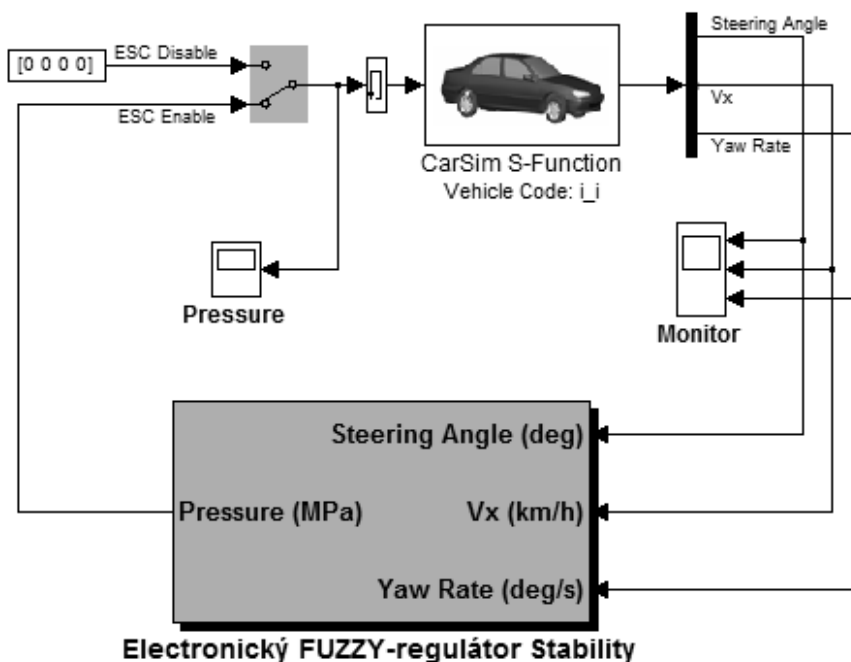
#### B. Návrh fuzzy riadenia stability

Pri návrhu fuzzy riadenia stability vozidla bol rovnako ako v pôvodnom ESP riadení využitý model vozidla a logika riadenia bŕzd, takže aj vstupy a výstup fuzzy regulátora sú rovnaké ako u pôvodného regulátora. V programe Matlab/Simulink bol vytvorený fuzzy regulátor typu Mamdani, pričom bolo použitých 25 rozhodovacích pravidiel podľa Tab.1.

Tabuľka 1  
 Rozhodovacie pravidlá pre fuzzy regulátor stability

de / e	ZV	Z	N	K	KV
ZV-záporná veľká	ZV	ZV	Z	Z	N
Z-záporná	ZV	Z	Z	N	K
N-nulová	Z	Z	N	K	K
K-kladná	Z	N	K	K	KV
KV-kladná veľká	N	K	K	KV	KV

Pri návrhu regulátora boli použité trojuholníkové a ako okrajové lichobežníkové funkcie príslušnosti. Na obr. 2 je uvedená schéma modelu vozidla s navrhnutým fuzzy regulátorom ESP v programe CarSim.



Obr. 2 Schéma riadenia stability vozidla s fuzzy regulátorom v programe CarSim.

### III. SIMULÁCIE A POROVNANIA NAVRHNUTÉHO RIEŠENIA

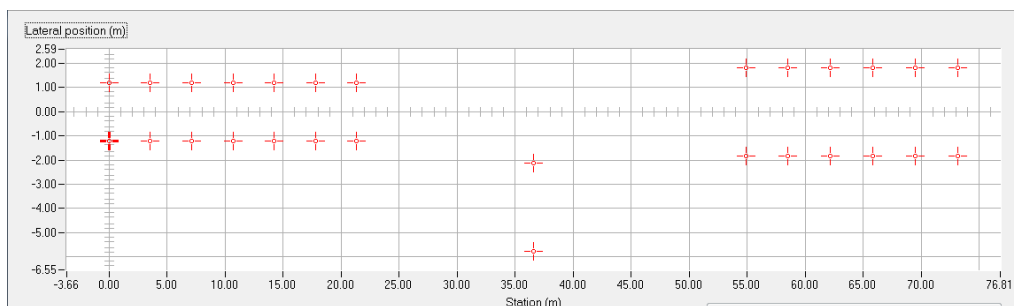
Na verifikáciu fuzzy riadeného ESP v programe CarSim boli použité simulácie jazdy v rôznych situáciách modelu vozidla vyššej strednej triedy s rovnakou konfiguráciou, pričom modré vozidlo bolo bez ESP regulátora, žlté bolo s klasickým a červené s fuzzy ESP regulátorom. Pre porovnanie sú v tomto článku z viacerých testovaných situácií vybrané dve testovania:

- pri obídení neočakávanej prekážky
- pri prejazde ostrou zákrutou

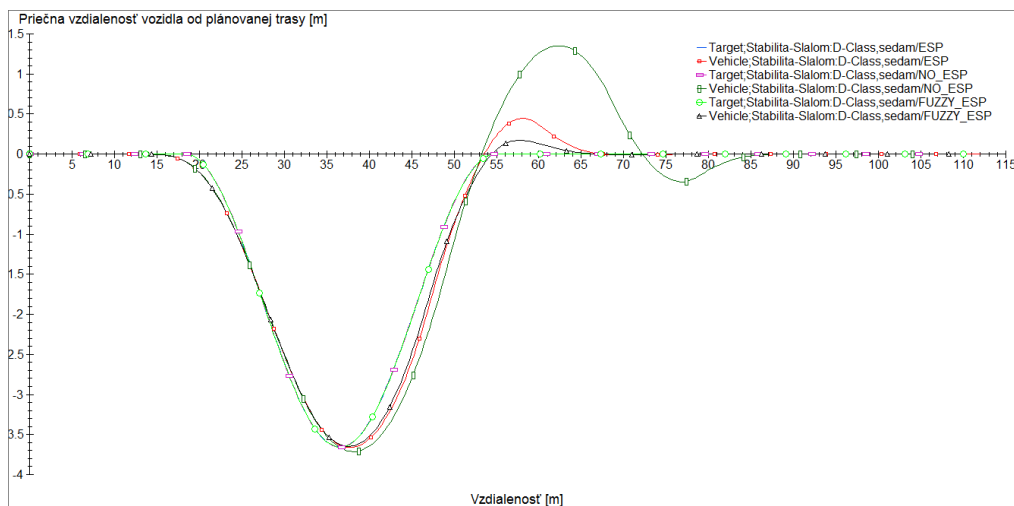
Všetky podmienky jednotlivých simulácií boli pre jednotlivé vozidlá rovnaké. Pri simuláciách vozidiel s ESP regulátorom a fuzzy ESP regulátorom bolo potrebné prepojenie programu CarSim s programom Simulink, v ktorom boli regulátory vytvorené.

#### A. Test fuzzy ESP pri nečakannej prekážke

Na obr. 3 je uvedený tvar testovacej dráhy, ktorá bola použitá pre simuláciu situácie, keď sa pred vozidlom objaví nečakaná prekážka, ktorú sa vodič snaží obísť, pretože by už nedokázal zastaviť vozidlo.



Obr. 3 Testovacia dráha s umiestnením kužeľov pre simuláciu prekážky



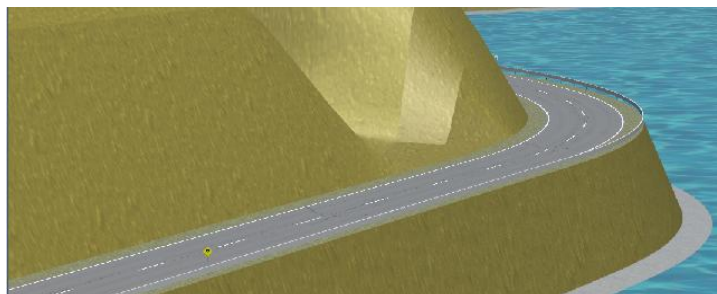
Obr. 4 Vychýlenie vozidiel od predpisanej stopy

Ako je zrejmé z obr.4, vozidlo s ESP (červený priebeh) má väčšie vychýlenie od požadovanej stopy (svetlozelený priebeh) než vozidlo s fuzzy ESP (čierny priebeh), pričom vozidlo bez ESP, ktoré nebolo pribrzďované, je v simuláciách uvedené zámerné pre porovnanie ako najhorší možný stav pri rovnakej pôvodnej rýchlosti vozidiel, ktorá bola v tomto teste 50km/h.

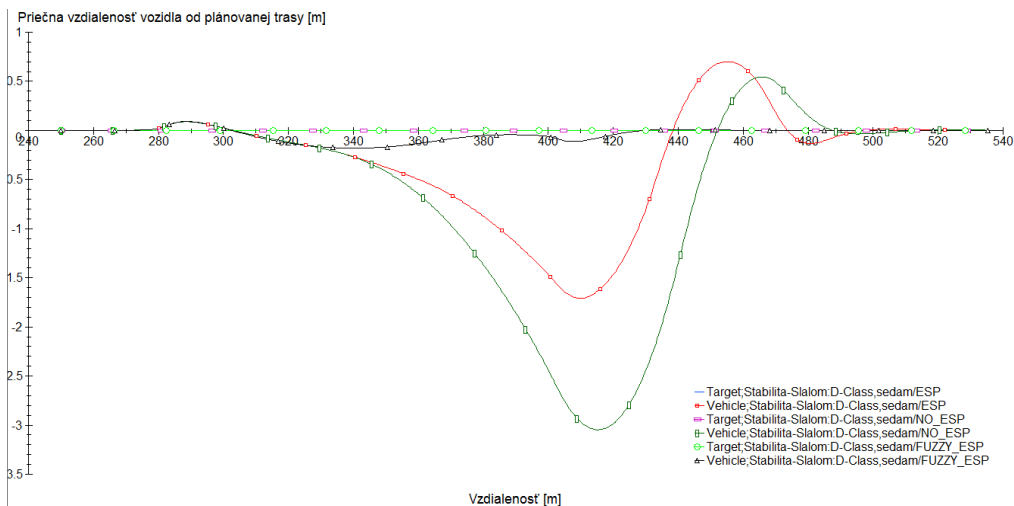
#### B. Test fuzzy ESP pri prejazde zákrutou

Pri teste fuzzy ESP pri prejazde zákrutou bola zvolená vozovka s ľavotočivou zákrutou s uhlom 90° (obr. 5), ktorej povrch tvorí ľad. Predpísaná stopa vozidla vedie stredom tejto vozovky a vodičom ovládané vstupy boli zvolené nasledovne:

- rýchlosť vozidla je konštantná 60 km/h
- vodič nebude brzdiť
- vozidlo pôjde stredom vozovky

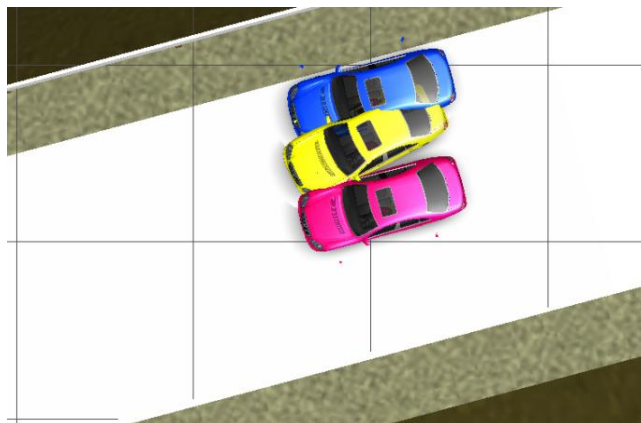


Obr. 5 Tvar vozovky pre simuláciu prejazd zákrutou



Obr. 6 Tvar vozovky pre simuláciu prejazd zákrutou

Výsledok testovania fuzzy ESP stabilizátora vozidla pri prejazde zákrutou je uvedený na obr.6. Podobne ako v predchádzajúcom teste, vozidlo s ESP (červený priebeh) má väčšie vychýlenie od požadovanej stopy ako vozidlo s fuzzy ESP (čierny priebeh).



Obr. 7 Poloha jednotlivých vozidiel po prejazde zákrutou

Výsledok animácie vozidiel po prejazde zákrutou uvedený na obr. 7 potvrdzuje, že vozidlo s fuzzy regulátorom ESP(červené) ide stredom vozovky, žlté vozidlo s ESP je vychýlené od predpísanej stopy a vozidlo bez ESP sa odchytilo od stredu vozovky až na jej krajinu.

V tejto simulácii je oproti iným simuláciám značný rozdiel medzi vozidlami s regulátorom ESP, čo je zapríčinené rýchlosťou, akou prechádzajú vozidlá túto zákrutu. Pri nižšej rýchlosti než bolo zvolených 60 km/h, by vo výsledkoch jednotlivých vozidiel boli menšie rozdiely.

#### IV. ZÁVER

Elektronické riadenie stability ESP prispieva k zníženiu dopravných nehôd, ktoré sú spôsobené hlavne v kritických situáciách, ako napríklad nezvládnutie zákruty a vybehnutie mimo vozovky alebo do protismeru, náraz do prekážky a pod. Zahraničné štúdie uvádzajú, že zavedenie ESP do nových vozidiel znižuje počet smrteľných dopravných nehôd samostatných vozidiel o 30 až 40 %, pričom u športových vozidiel to je až 67 % [1]. Tento druh dopravných nehôd sa najčastejšie stáva na diaľniciach, často pri zmenách počasia, pri riadení na vozovkách s mokrým povrchom, pri poľadovici a pod. Túto technológiu výrobcovia automobilov používajú vo svojich vozidlách pod rôznymi názvami, pričom rozdiel je zvyčajne v rýchlosti odozvy systému. Očakáva sa, že všetky novoregistrované automobily od novembra 2014 musia mať elektronický stabilizačný program ESP a uvažuje sa aj o zjednotení jeho názvu, ktorý by mal používať označenie ESC (Electronic Stability Control) [1].

Tento článok na príklade vozidla s jednoduchým ESP regulátorom a fuzzy ESP regulátorom, ktorý dosiahol ešte lepšie výsledky, je príkladom ako elektronický regulátor stability pomáha vodičovi prekonať krízové situácie pri nepriaznivých podmienkach a zvyšuje bezpečnosť prevádzky vozidla v cestnej premávke. Doplnením tohto systému o ďalšie prvky aktívnej bezpečnosti, ako je ABS, ASR, BAS a ďalšími, sa bezpečnosť prevádzky vozidla zvyšuje ešte viac.

Z porovnania testovaných modelov vozidla vyplýva, že situácia, ktorú zvládne vozidlo s niektorým z ESP regulátorov, je pre vozidlo bez ESP kritická a veľmi nebezpečná a v bežnej cestnej prevádzke takémuto vozidlu hrozí reálne kolízia.

#### POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.

#### POUŽITÁ LITERATÚRA

- [1] ESP je jednoznačný pomocník[Online] <http://www.becep.sk/vodici/>
- [2] CarSim, [Online] Dostupné na internete: <http://www.carsim.com/>
- [3] F. Vlk, "Elektronické systémy motorových vozidiel 2", Brno, 2002. s. 299-359
- [4] F. Vlk, "Dynamika motorových vozidiel", Brno, 2003. s. 265-432
- [5] Bezpečné vozidlo,[Online] Dostupné na internete: <http://www.ibesip.cz/>
- [6] Slovenská správa ciest: Bezpečnosť cestnej premávky. [online] Dostupné na internete: <[http://www.ssc.sk/files/documents/rozvoj\\_cestnej\\_siete/koncepcia/04\\_bezpecnost.pdf](http://www.ssc.sk/files/documents/rozvoj_cestnej_siete/koncepcia/04_bezpecnost.pdf)>

# Riadenie veternej elektrárne

<sup>1</sup>Peter GIROVSKÝ, <sup>2</sup>Ján KAŇUCH, <sup>3</sup>Milan LACKO, <sup>4</sup>Miloslav URAM

<sup>1,2,3</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

<sup>1</sup>peter.girovsky@tuke.sk, <sup>2</sup>jan.kanuch@tuke.sk, <sup>3</sup>milan.lacko@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá návrhom riadenia natočenia veternej elektrárne v závislosti od veterných podmienok. Návrh riadenia bol robený pre fyzikálny model veternej elektrárne, na ktorom bolo navrhnuté riadenie aj overené. V tomto príspevku bude predstavené konkrétne riešenie natáčania veternej elektrárne v závislosti od smeru a rýchlosti vetra.

**Kľúčové slová** — veterná elektráreň, riadenie, anemometer, snímač

## I. ÚVOD

V súčasnej dobe je čoraz väčší dopyt po elektrickej energii. Väčšina elektrickej energie sa vyrába z neobnoviteľných zdrojov, akými sú ropa, plyn či uhlie a len malá časť využíva obnoviteľné zdroje. Jeden z najrozšírenejších druhov takzvanej zelenej energie je vietor. Energia vetra sa využíva už od 18. Storočia, najskôr sa premieňala iba na mechanickú energiu, neskôr aj na elektrickú energiu.

Pre efektívne využitie sily vetra je potrebné vybrať nielen správne umiestnenie elektrárne, ale aj rozmery, tvar lopatiek a výšku veže. No aj po splnení všetkých týchto podmienok je smer a rýchlosť vetra premenlivá veličina, čo spôsobuje nestálosť výstupného výkonu. Aby sme zabezpečili čo najefektívnejší, plynulý a bezpečný chod je potrebné veternú elektráreň riadiť. Zber potrebných údajov zabezpečuje anemometer, po ich spracovaní riadiaca jednotka vhodným zásahom zmení parametre elektrárne tak, aby využila čo najväčšie množstvo veternej energie. V prípade hraničných podmienok, akými sú napríklad prekročenie konštrukčnej rýchlosti či porucha, musí byť riadenie schopné veternú elektráreň bezpečne zastaviť.

Veterné turbíny sú navrhnuté tak, aby dodávali čo najväčší výkon pri rôznej rýchlosti vetra. Pre každú veternú turbínu je daná nominálna rýchlosť, pri ktorej turbína dosahuje nominálny výkon a maximálna rýchlosť, ktorej prekročením dochádza k deštrukcii turbíny. Riadenie veternej turbíny je závislé od rýchlosti vetra a podľa toho rozoznávame tri základné prevádzkové režimy veterných turbín:

- režim pod nominálnou rýchlosťou vetra,
- režim v okolí nominálnej rýchlosti vetra,
- režim nad nominálnou rýchlosťou vetra.

## II. NÁVRH RIADENIA VETERNEJ ELEKTRÁRNE

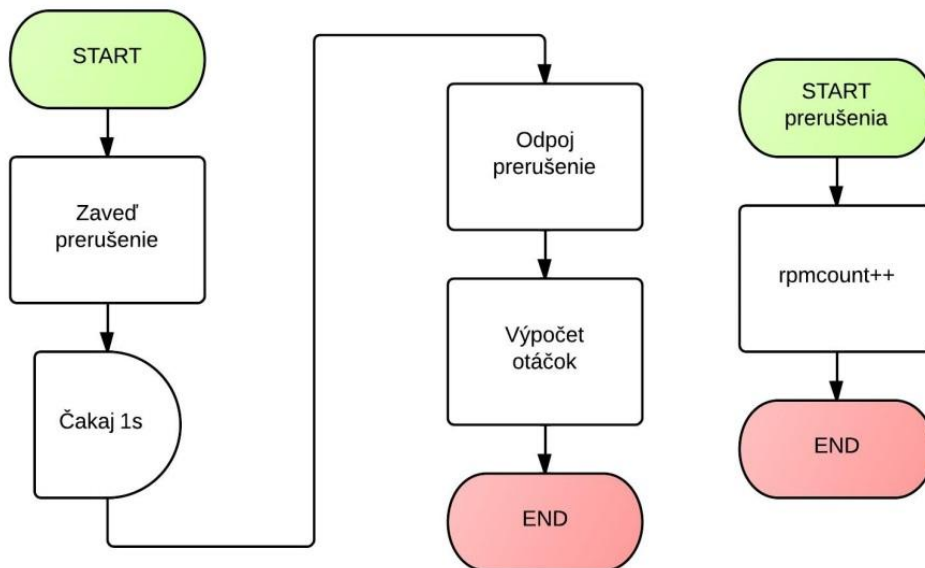
Pre návrh riadenia veternej elektrárne potrebujeme poznať základné údaje o smere a rýchlosti vetra. Na základe poznania týchto dvoch veličín sa bude riadený model veternej elektrárne. Na získanie týchto informácií nám slúžil anemometer, ktorý pozostával zo snímača rýchlosti a snímača polohy, resp. natočenia.

### A. Meranie rýchlosti vetra

Pre meranie rýchlosti vetra bol použitý snímač, ktorého základ bol tvorený vysielačom, prijímačom a diskom so zárezmi po obvode. Disk bol pripevnený na zvislej rotujúcej osi, na ktorej vrchole nachádzali lopatky. Prerušovaním signálu vznikajú impulzy, ktoré sú posielané riadiacej jednotke na ďalšie spracovanie. Počet vygenerovaných impulzov je priamo úmerný rýchlosti vetra.

Ako vysielač slúži infračervená dióda a prijímač je fototranzistor, ktorého signál je privedený do mikroprocesora na spracovanie. V rámci riadiaceho programu sa počíta počet

dobehových hrán za jednu sekundu a na základe tejto hodnoty sa vypočítavajú otáčky a z nich sa určí rýchlosť vetra. Princiálny vývojový diagram snímania otáčok (rýchlosti vetra) je zobrazený na obr.1.



Obr. 1 Vývojový diagram snímania otáčok (rýchlosti vetra).

### B. Snímanie smeru vetra

Základom pre snímanie smeru vetra je snímač tvorený štyrmi optickými závorami a kódovacím diskom. Optické závory sú tvorené infračervenými diódami a fototranzistormi a sú umiestnené v jednej línii tak, aby bol signál prerušovaný výrezmi v kódovacom disku. Kódovací disk je tvorený priehľadným plexisklom, na ktorom je umiestnená maska, ktorá zabezpečuje prerušovanie signálov.

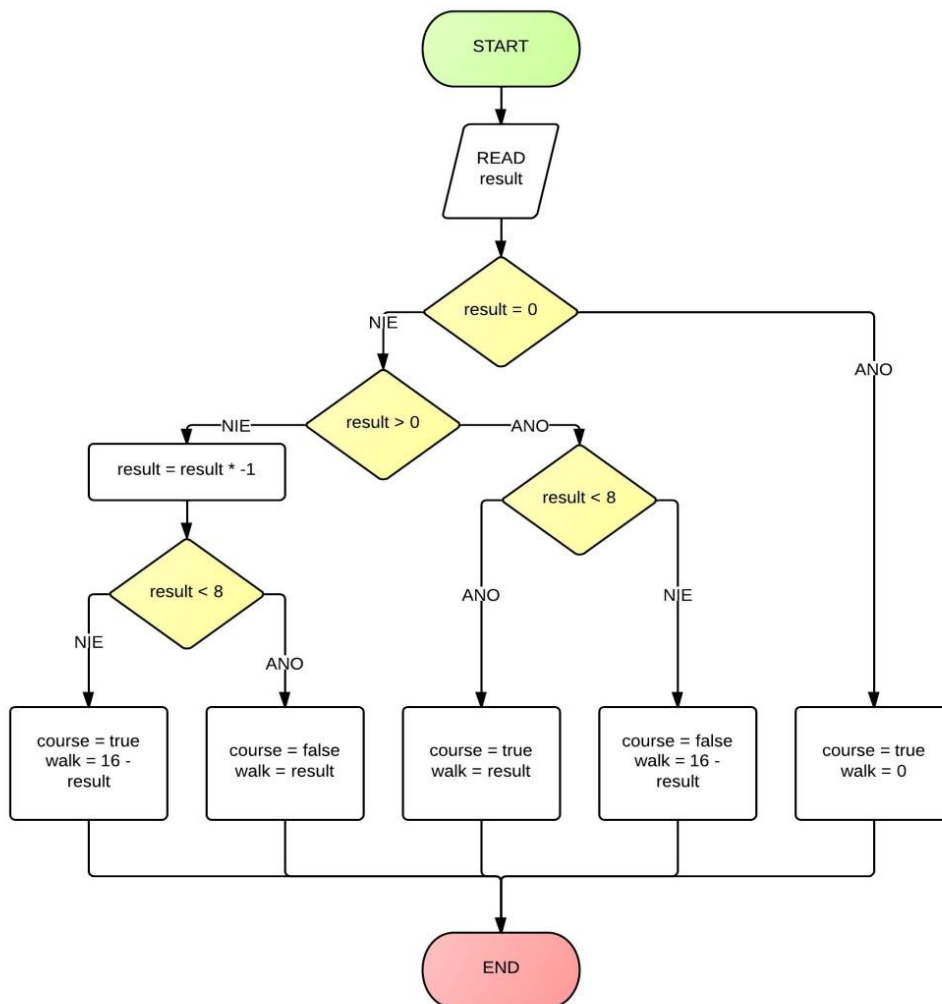
Disk je pripevnený na rotujúcej osi. Na nej je pripevnená aj smerová lopatka. Ak začne sila vetra pôsobiť na lopatku, vychýli ju z aktuálnej polohy, čo spôsobí otočenie disku a tým sa zmení binárny kód. Celý kruh (360°) je rozdelený na 16 častí (0-15), z ktorých každá má jedinečný 4-bitový kód. Každá z týchto častí má 22,5°, z čoho vyplýva, že tento snímač dokáže určiť smer vetra s presnosťou +/- 11,25°. Výstupom tohto snímača je 4-bitová informácia, ktorá sa ďalej spracúva ako binárne číslo podľa tab.1.

Tabuľka 1  
 Tabuľka stavov snímača smeru vetra.

STAV	A	B	C	D
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1

<b>12</b>	1	1	0	0
<b>13</b>	1	1	0	1
<b>14</b>	1	1	1	0
<b>15</b>	1	1	1	1

Principiálny vývojový diagram snímania smeru vetra je zobrazený na obr.2.



Obr. 2 Vývojový diagram snímania smeru vetra.

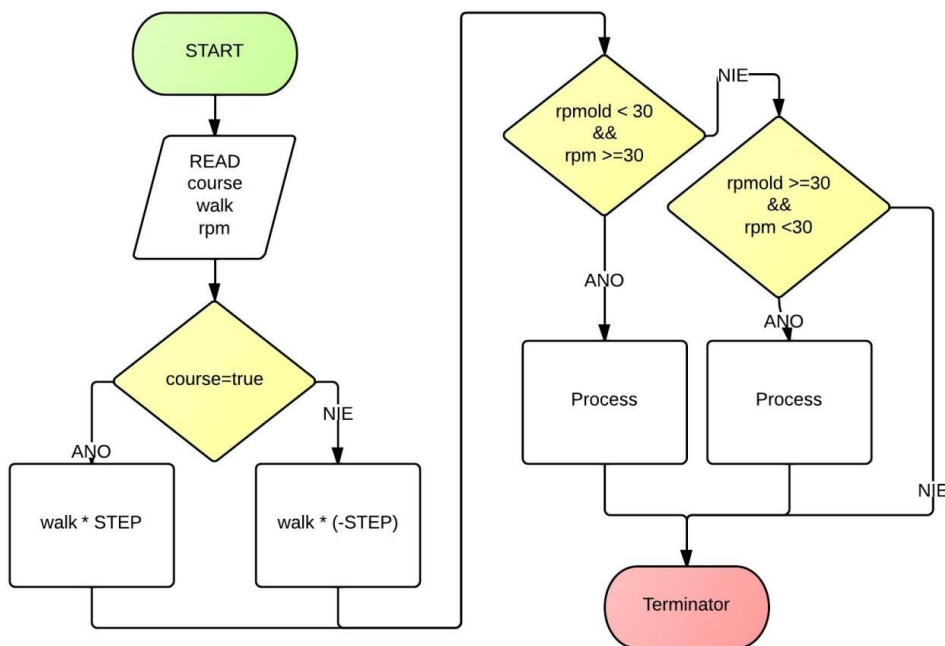
### C. Riadenie natáčania veternej turbíny

Pred začiatkom spustenia riadenia natáčania veternej elektrárne je potrebné najprv spustiť inicializačnú procedúru, ktorá zabezpečí počiatočné podmienky pri riadení. Počiatočnou podmienkou je, že turbína aj snímač smeru vetra musia byť v rovnakej, nulovej polohe. Ak je táto podmienka splnená, môže sa spustiť hlavný riadiaci program pre riadenie natáčania veternej elektrárne. Vývojový diagram hlavného riadiaceho programu je zobrazený na obr.3.

Hlavný riadiaci program získava informácie od snímača rýchlosti a smeru vetra a po vyhodnotení údajov z týchto snímačov zabezpečí prostredníctvom krokového motora správne natočenie veternej turbíny.

Riadiaci program najprv určí smer natočenia podľa hodnoty príslušnej premennej od snímača smeru vetra. Po natočení turbíny na správny smer overí či nebola prekročená kritická rýchlosť vetra, a ak sa táto rýchlosť prekročí, riadenie nezávisle na polohe natočí turbínu o 512 krokov, čo predstavuje uhol 90°, čím sa zabezpečí zastavenie veternej turbíny.





Obr. 3 Vývojový diagram hlavného riadiaceho programu.

### III. ZÁVER

V dnešnej dobe je problematika veterných elektrární ako aj ostatných obnoviteľných zdrojov energie vysoko aktuálna. Cieľom tohto príspevku bolo oboznámiť čitateľov s jednoduchým návrhom riadenia fyzikálneho modelu veternej elektrárne.

Základom modelu veternej elektrárne bol anemometer, ktorý pomocou optických snímačov vyhodnocoval smer a rýchlosť vetra. Získané údaje odovzdá riadiacej jednotke, ktorá ich podľa riadiaceho programu vyhodnotí a v prípade potreby dokáže riadiaca jednotka vykonať príslušný zásah a prostredníctvom krokového motora natočiť maketu veternej elektrárne tak, aby vždy boli splnené podmienky kladené na bezpečnosť a efektívnosť využívania veternej energie.

### POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.

Táto práca bola podporená projektom KEGA - 042TUKE-4/2012.

### LITERATÚRA

- [1] P. Gipe, Wind Power: Renewable Energy for Home, Farm, and Buisness, 2nd edition. White River Junction VT: Chelsea Green Pub. Co., 1993. 512. ISBN 978-1-931498-14-2
- [2] E. Hau, Wind Turbines. Springer, 2006. 783. ISBN 978-3-540-24240-6
- [3] P. Brondsted, Advances in Wind Turbine Blade Design and Materials. Woodhead Pub., 2013. 464. ISBN 978-0-85709-426-1
- [4] V. Fred, Meteorological Measurement Systems. Oxford university Press., 2001. 304. ISBN 0-19-513451-6

# Simulácia ASR systému

<sup>1</sup>Peter GIROVSKÝ, <sup>2</sup>Jaroslava ŽILKOVÁ, <sup>3</sup>Ervín BERNÁT

<sup>1,2</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

<sup>1</sup>peter.girovsky@tuke.sk, <sup>2</sup>jaroslava.zilkova@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá návrhom a realizáciou simulácie ASR systému vozidla, pričom boli uvažované rôzne prístupy k riadeniu ASR systému. Prvý prístup vychádzal z využitia regulátora bŕzd, druhý z využitia regulátora škrtiacej klapky a tretí z kombinácie prvých dvoch prístupov. Návrh a simulácia asistenčného systému ASR pre vozidlo bolo realizované za pomoci programov MATLAB/Simulink a CarSim. V závere príspevku je uvedená simulačná verifikácia správania sa vozidla bez systému ASR a so systémom ASR na kombinovanej vozovke.

**Kľúčové slová** — asistenčné systémy vozidla, ASR (Systém pre reguláciu preklzovania kolies), CarSim

## I. ÚVOD

V dnešnej dobe je vozidlo neoddeliteľnou súčasťou každého života človeka, či už v práci alebo v súkromnom živote. Počty vozidiel na cestách sa každým rokom zvyšujú a so zvyšujúcim sa počtom vozidiel na cestách sa aj zvyšuje riziko nehôd. V moderných vozidlách sa riziko nehôd potláča zavádzaním rôznych bezpečnostných systémov či rôznych pomocných asistentov, ktorých úlohou je buď varovať vodiča pred blížiacim sa nebezpečenstvom, alebo pomáhať mu pri rôznych hraničných situáciách, pri brzdení vozidla alebo pri jeho rozbehu.

Bezpečné rozbehy vozidiel na cestách tvoria základ pre bezpečné vedenie vozidla v cestnej premávke. Mnohí vodiči majú problém sa rozbehnúť s vozidlom na vozovke za zhoršených adhézných podmienkach. Vďaka pomocným asistenčným systémom ako je napríklad ASR môžeme riziko problémov pri rozbehu eliminovať.

Protišmykový systém ASR predstavuje systém regulácie prešmyku kolesa a prvýkrát bol použitý vo vozidle už v roku 1986. Úlohou tohto systému je zaistenie stability a ovládateľnosti vozidla pri jeho zrýchľovaní. Výhodou systému ASR, okrem toho, že zabraňuje poháňanému kolesu v prešmyku je aj nasledovné:

- zabezpečuje stabilné jazdné vlastnosti pri rozjazde, zrýchľovaní a prejazde áut na klzkých vozovkách,
- upozorňuje vodiča na klzké vlastnosti vozovky,
- znižuje opotrebovanie pneumatík na minimum,
- chráni pohonné ústrojenstvo vozidla,
- znižuje riziko nehody.

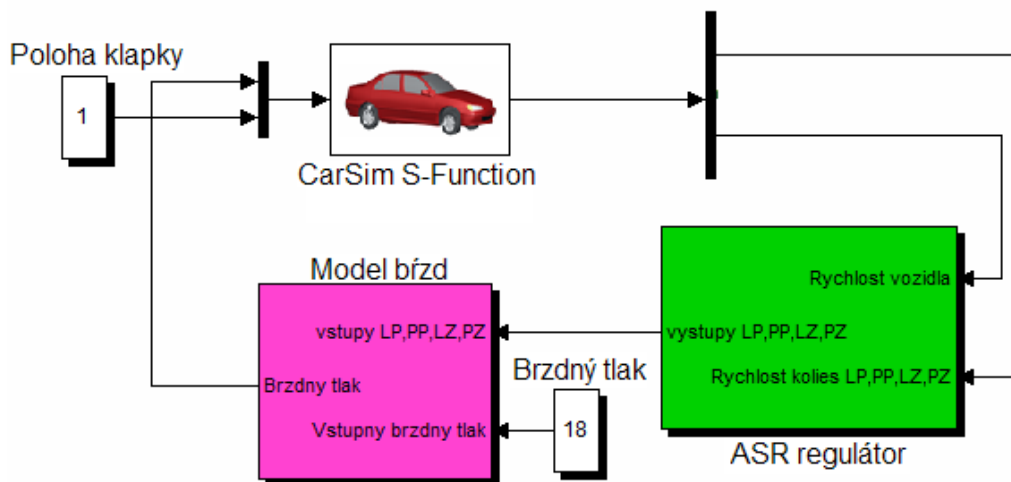
## II. NÁVRH SYSTÉMU ASR

Pri návrhu ASR systému vozidla sme uvažovali s troma prístupmi pre obmedzenie hnacieho momentu a tým aj zamedzeniu prešmyku hnacieho kolesa pri akcelerácii vozidla:

- regulácia ASR pomocou brzdneho tlaku
- regulácia ASR pomocou škrtiacej klapky
- regulácia ASR pomocou brzdneho tlaku a škrtiacej klapky

### A. Regulácia ASR pomocou brzdneho tlaku

Pri tomto spôsobe riadenia ASR systému celkovú reguláciu vykonáva brzdny systém vozidla a poloha škrtiacej klapky je nastavená na hodnotu 1, čo znamená, že výkon motora sa nereguluje. Schéma regulácie ASR systému pomocou brzdneho tlaku je znázornená na obr.1.

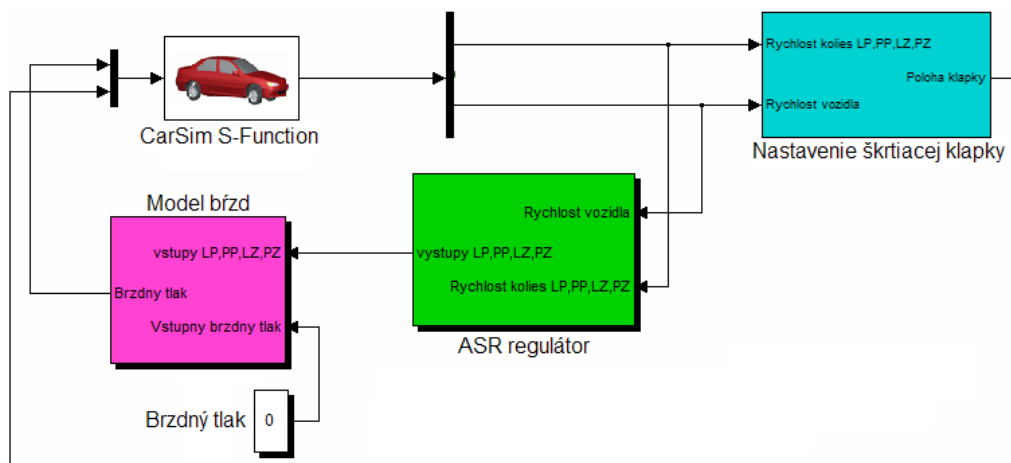


Obr. 1 Regulácia ASR systému pomocou brzdného tlaku.

Princíp tejto regulácie spočíva v identifikácii prešmyku kolesa, pri ktorom sa porovnáva rýchlosť vozidla s rýchlosťou jednotlivých kolies, pričom každé z kolies sa reguluje samostatne. Na výstupe z regulátora môžu byť nastavené dve hodnoty, 0 alebo 1. Na začiatku je nastavená hodnota výstupu 0, ktorá je nastavená pokiaľ prešmyk neprekročí 10%. V okamihu ak prešmyk dosiahne hodnotu 10%, sa na výstupe nastaví hodnota 1, čo znamená, že vozidlo začne pribrzďovať potrebné koleso. Ak hodnota sklzu klesne pod úroveň 5% výstup opäť nadobudne hodnotu 0, vozidlo prestane pribrzďovať a opäť akceleruje.

#### B. Regulácia ASR pomocou škrtiacej klapky

Pri tomto spôsobe riadenia ASR sa regulácia vykonáva iba riadením polohy škrtiacej klapky, ktorá obmedzuje prísun vzduchu do spaľovacieho motora. Schéma regulácie ASR systému pomocou škrtiacej klapky je znázornená na obr.2.

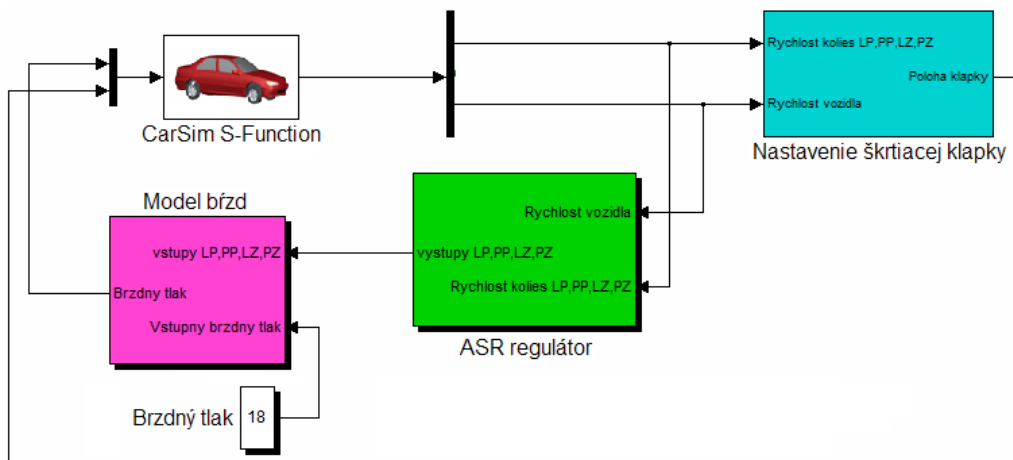


Obr. 2 Regulácia ASR systému pomocou škrtiacej klapky.

Princíp činnosti pri tomto type regulácie vychádza taktiež z výpočtu prešmyku ako v predchádzajúcom prípade. Vynechanie bŕzd je v tomto prípade simulované nastavením vstupného brzdného tlaku na hodnotu 0. Výstup z nastavenia škrtiacej klapky nadobúda dve stavy 0, keď je škrtiaca klapka zatvorená a 1, keď je škrtiaca klapka otvorená. Na začiatku je škrtiaca klapka otvorená a ak sa ľubovoľné koleso hnacej nápravy dostane do prešmyku 10%, klapka sa zatvorí. Po poklese prešmyku na hodnotu 5% sa klapka opätovne otvorí a do motora je opäť dodávaný vzduch. Pre výpočet prešmyku sa berie do úvahy vždy najvyššia rýchlosť ľubovoľného kolesa.

### C. Regulácia ASR kombináciou brzd a škrtiacej klapky

Pri tomto spôsobe riadenia ASR sa regulácia vykonáva kombináciou regulácie výkonu motora pomocou škrtiacej klapky a príbrzdovaním kolies. Schéma tohto kombinovaného spôsobu regulácie ASR systému je znázornená na obr.3.



Obr. 3 Regulácia ASR systému kombináciou brzd a škrtiacej klapky.

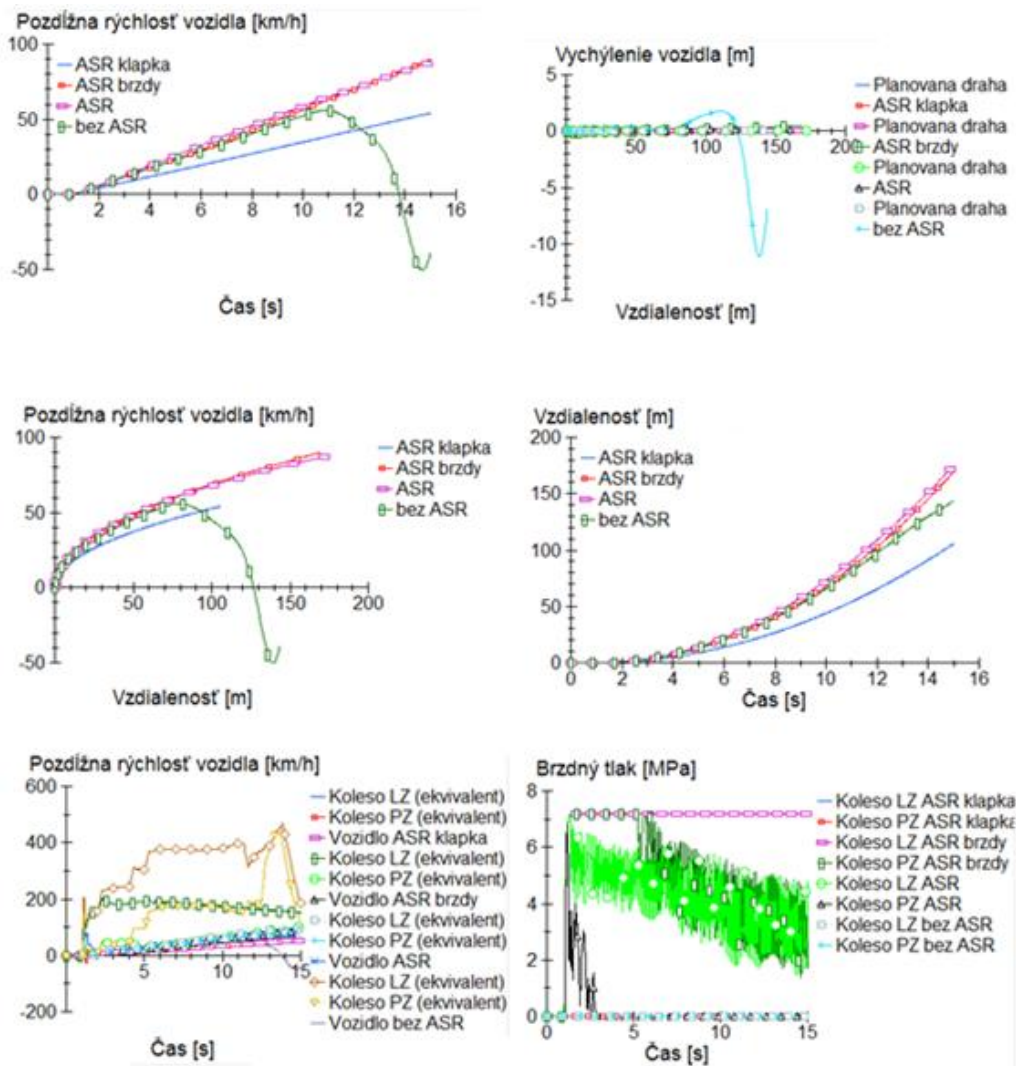
Princíp činnosti pri tomto type regulácie vychádza taktiež z výpočtu prešmyku ako v predchádzajúcich dvoch prípadoch, pričom pri prešmyku hnacej nápravy dochádza k súčasnému zníženiu výkonu motora pomocou škrtiacej klapky, ako aj príbrzdovaniu jednotlivých kolies. Výsledky testu rozbehu na danej vozovke sú zobrazené na obr. 4 a v tab.1.

### III. TESTOVANIE ASR REGULÁTORA

Simulačné testovanie navrhnutých systémov riadenia ASR bolo robené pomocou simulačných programov MATLAB/Simulink a CarSim. Test vozidla s jednotlivými systémami ASR a vozidla bez ASR bol vykonaný na rovnej ceste s kombinovaným povrchom, kde ľavý pruh vozovky mal priľnavosť 0,2 čo predstavovalo ľad na ceste a pravý pruh vozovky mal priľnavosť 0,5 čo predstavuje klasický asfalt. Testované vozidlo bolo umiestnené v strede vozovky. Výsledky testu rozbehu na danej vozovke sú zobrazené v tab.1 a na obr. 4.

Tabuľka 1  
 Výsledky testov vozidiel s jednotlivými systémami ASR a bez systému ASR na kombinovanej vozovke.

		Bez ASR	ASR (brzdy)	ASR (klapka)	ASR (klapka + brzdy)
Vzdialenosť	[m]	143,79	169,527	105,748	175,14
Maximálna rýchlosť	[km/h]	55,948	89,89	54,083	87,707
Maximálna rýchlosť L kolesa	[km/h]	470,79	201,34	172,763	138,49
Maximálna rýchlosť P kolesa	[km/h]	442,57	128,83	117,987	89,49
Konečná rýchlosť	[km/h]	-38,33	89,89	54,083	87,707
Počiatková rýchlosť	[km/h]	0	0	0	0
Maximálne vychýlenie P	[m]	-11,11	-0,042	-0,0019	-0,009
Maximálne vychýlenie L	[m]	1,748	0,419	0,011	0,155
Konečná vychýlenie	[m]	-6,897	0,411	0,011	0,155



Obr. 4 Priebehy testov vozidiel s jednotlivými systémami ASR a bez systému ASR na kombinovanej vozovke.

Z testu vyplýva, že vozidlo s kombinovaným systémom ASR, kombinácia riadenia polohy škrtiacej klapky a príbrzdovana prešmykujúceho sa kolesa, je najvhodnejší. Vozidlo s takýmto systémom prejde najväčšiu vzdialenosť 175,14 m a má malé vybočenie z plánovanej dráhy 15,5 cm. Vozidlo bez systému ASR je na kombinovanom povrchu nestabilné, vozidlo dostane šmyk vo vzdialenosti približne 74 metrov. Na konci testu sa toto vozidlo nachádza mimo vozovky.

Využitie systému ASR s reguláciou polohy škrtiacej klapky je na vychýlenie od plánovanej dráhy najpresnejšie, kde vychýlka je iba 1,1 cm, ale vozidlo má približne 62 % rýchlosť vozidla s kombinovaným systémom ASR. Systém ASR s reguláciou brzd dokáže vyvinúť veľkú rýchlosť vozidla ale má pomerne veľké vychýlenie 41,1 cm.

#### IV. ZÁVER

V tomto príspevku bol prezentovaný návrh systému ASR v programe MATLAB/Simulink a CarSim s tromi typmi regulácie, pomocou brzd, škrtiacej klapky a ich kombináciou. V následných testoch navrhnutého riadenia ASR systému sme sledovali vychýlenie vozidla, prejdenná vzdialenosť, brzdňý tlak pôsobiaci na kolesá a iné.

Z výsledkov týchto testov na kombinovanej vozovke vyplýva, že vozidlo využívajúce systém ASR s kombináciou regulácie brzdňého tlaku a polohy škrtiacej klapky dosahovalo najlepšie výsledky a tento preto je tento systém najvhodnejší pre použitie vo vozidlách.

Ďalšími úpravami by bolo možné dosiahnuť ešte lepšie výsledky pre jednotlivé spôsoby regulácie, prípadne pre riadenie ASR využiť systém založený na fuzzy logike, prípadne neurónových sietach.

#### POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.



Podporujeme výskumné aktivity na Slovensku/ Projekt je spolufinancovaný zo zdrojov EÚ. Tento článok bol vypracovaný v rámci projektu "Centrum excelentnosti integrovaného výskumu a využitia progresívnych materiálov a technológií v oblasti automobilovej elektroniky", ITMS 26220120055.

#### LITERATÚRA

- [1] L. Miková, R. Baláž, M. Ádiová, Mechatronické subsystémy vozidiel pre zlepšenie jazdných vlastností a bezpečnosti cestných vozidiel. In: AT&P journal PLUS. 2009, č. 1, s. 13-17.
- [2] A. Mareš, AKTÍVNE A PASÍVNE PRVKY BEZPEČNOSTI AUTOMOBILOV. Archív: Transfer inovácií. [Online] 2003. <http://www.sjf.tuke.sk/transferinovacii/pages/archiv/transfer/6-2003/pdf/193-194.pdf>. ISBN 80 – 8075 – 075 – X.
- [3] WABCO: Anti-Lock Braking System (ABS) and Anti-Slip Regulation (ASR) [Online]. 2011. 02.2011. Dostupné na internete: <<http://inform.wabco-auto.com/intl/pdf/815/01/94/8150101943.pdf>>.

# Využitie vizualizácie pri výučbe algoritmov a údajových štruktúr

Slavomír ŠIMONÁK

Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika

slavomir.simonak@tuke.sk

**Abstrakt** – Témou príspevku je ozrejenie motivácie pre využitie vizualizácie v rámci výučby algoritmov a údajových štruktúr, prezentácia aktuálneho stavu vizualizačnej platformy VizAlgo, ako aj načrtnutie niektorých možností pre pokračovanie výskumu a vývoja v danej oblasti v našich podmienkach.

**Kľúčové slová** – vizualizácia algoritmov a údajových štruktúr, platforma VizAlgo, zásuvné moduly

## I. ÚVOD

Poznatky z oblasti návrhu a analýzy algoritmov a údajových štruktúr majú pevné miesto v osnovách študijných programov vysokoškolského štúdia zameraného na oblasť informatiky a informačných technológií [1], [8]. Za jednu zo všeobecne akceptovaných [3], [12] možností zefektívnenia [7] procesu výučby v tejto oblasti je považované využitie vizualizácií [2]. V porovnaní s inými metódami pre podporu výučby algoritmov (ako napr. pseudokódy, diagramy), prednosťou vizualizácií je zachytenie dynamiky vykonávania algoritmu a transformácie údajových objektov s ktorými algoritmus pracuje. Pri výučbe algoritmov a údajových štruktúr v rámci predmetu Údajové štruktúry a algoritmy, vyučovaného v bakalárskom stupni študijného programu Informatika, využívame kombináciu viacerých podporných metód, pričom využitie vizualizácie považujeme v tejto oblasti za veľmi perspektívne [11]. Určitým problémom pri využití vizualizácie je dostupnosť vhodných softvérových podporných nástrojov. Hoci takých nástrojov je v súčasnosti dostupných pomerne veľa [5], nie je ľahké nájsť jeden, ktorý plne vyhovuje potrebám pre nasadenie napr. v rámci konkrétneho predmetu. Preto sme sa po analýze možností [10] rozhodli vytvoriť novú platformu pre vizualizáciu algoritmov a údajových štruktúr, ktorá by čo najlepšie vyhovovala našim požiadavkám.

## II. PLATFORMA VIZALGO

Vytvorenou platformou je VizAlgo, pričom medzi jej základné vlastnosti, na ktoré sme brali ohľad pri návrhu a vývoji patrí prenositeľnosť a možnosť jednoduchého rozšírenia funkcionality. Prenositeľnosť je zabezpečená voľbou technológie Java, ktorej vykonávacie prostredie je dostupné v súčasnosti na mnohých softvérových i hardvérových platformách. Rozširovanie aplikácie o nové vizualizácie je zabezpečené pomocou zásuvných modulov. Pre podporu a zjednodušenie vývoja zásuvných modulov bol zvolený rámec JSPF [4]. Detailný opis architektúry platformy VizAlgo, ako aj spôsob tvorby zásuvných modulov je možné nájsť v [9] a [6]. Používateľské prostredie aplikácie je možné vidieť na obrázku Obr. 5, jeho detailnejší opis je dostupný napr. v [10].

## III. DOSTUPNÉ VIZUALIZÁCIE

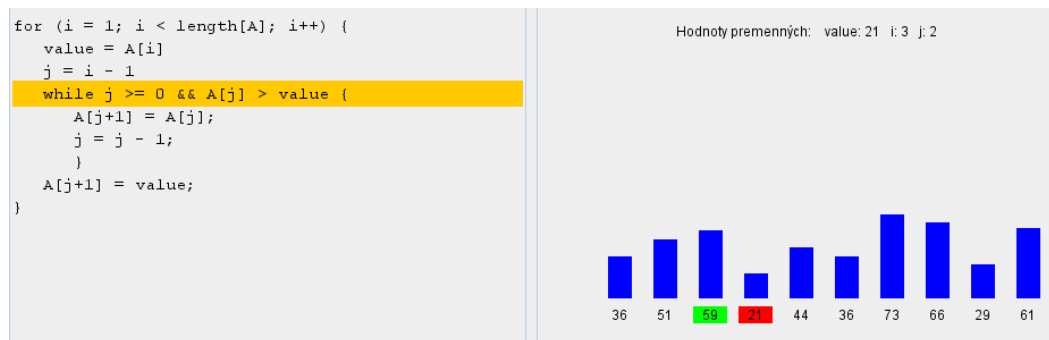
V súčasnosti dostupné vizualizácie vo forme zásuvných modulov sme rozdelili do niekoľkých skupín (algoritmy triedenia, dynamické programovanie, rozdeľuj a panuj, rekurzívne algoritmy, vizualizácia údajových štruktúr) podľa oblasti určenia, resp. podľa použitej metódy návrhu jednotlivých algoritmov. Je preto pochopiteľné, že niektoré algoritmy by mohli byť zaradené aj do viacerých oblastí súčasne (napr. *Quicksort*, ktorý je algoritmom triedenia, ale zároveň využíva techniky dekompozície a rekurzie).

Z priestorových dôvodov nie je v tomto príspevku možný detailný opis činnosti všetkých doposiaľ vytvorených zásuvných modulov. Pokúsime sa preto vybrať niekoľko z nášho pohľadu zaujímavých, typovo odlišných vizualizácií a tak priblížiť čitateľovi platformu VizAlgo formou ich opisu. Pri prezentácii aplikácie zameranej na vizualizáciu považujeme za výhodné textový opis doplniť vo zvýšenej miere názornými ilustráciami.



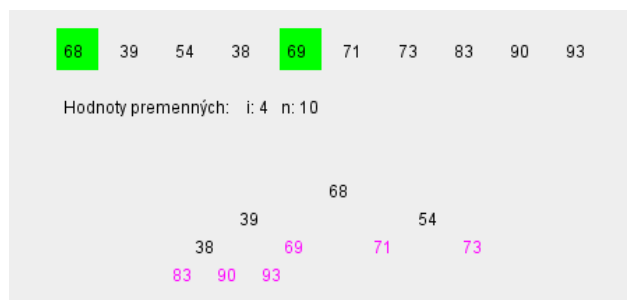
### A. Algoritmy triedenia

Z jednoduchších algoritmov triedenia založených na vzájomnom porovnaní a výmene triedených prvkov boli vytvorené vizualizácie algoritmov *Bubblesort* (bublinové triedenie), *Insertsort* (triedenie vkladáním), *Selectsort* (triedenie výberom), *Shellsort* (Shellovo triedenie). Implementačne sa jedná o jednoduchšie algoritmy a ich asymptotická zložitosť sa pohybuje na úrovni  $O(n^2)$ .



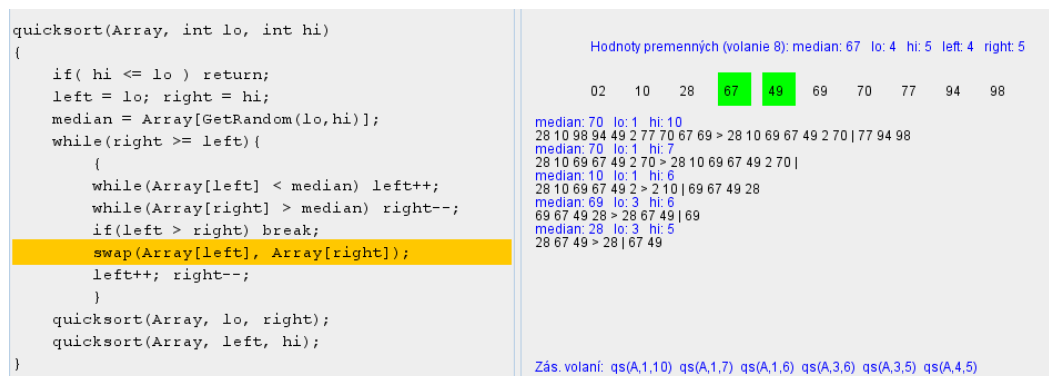
Obr. 1 Triedenie vkladáním (Insertsort).

Z vyššie uvedených algoritmov je pre ilustráciu zobrazená vizualizácia algoritmus Insertsort (Obr. 1). V ľavej časti sa nachádza pseudokód algoritmu so zvýrazneným aktuálne vykonávaným riadkom, v pravej časti sa nachádza vizualizácia činnosti algoritmu na konkrétnej postupnosti údajov. Použité je číselné a pre zvýšenie názornosti aj stĺpcové zobrazenie triedených hodnôt, ako aj aktuálne hodnoty premenných. Farebne sú zvýraznené aktuálne porovnávané prvky.



Obr. 2 Triedenie haldou (Heapsort).

Z efektívnejších triediacich algoritmov sú dostupné vizualizácie algoritmov *Heapsort* (triedenie haldou), *Quicksort* (rýchle triedenie), *Radixsort* (radixové triedenie) a *Mergesort* (triedenie zlúčovaním). V rámci vizualizácie algoritmu *Heapsort* (Obr. 2) s časovou zložitosťou  $O(n \cdot \log n)$  je okrem postupnosti triedených prvkov zobrazený aj binárny strom reprezentujúci haldu. Prvky v strome aktuálne spracovávané algoritmom sú zobrazené čiernou farbou, zatiaľ čo prvky už umiestnené na správnych pozíciách sú zobrazené odlišnou farbou.



Obr. 3 Vizualizácia algoritmu Quicksort.

Ďalším z efektívnych triediacich algoritmov, pre ktoré je dostupná vizualizácia v rámci platformy



VizAlgo je *Quicksort* (Obr. 3). Okrem vizualizácie pseudokódu, aktuálnych hodnôt premenných a postupnosti triedených prvkov je zobrazovaný aj zásobník volaní, nakoľko sa jedná o rekurzívny algoritmus. Položky v tomto zásobníku volaní reprezentujú zásobníkové rámce pre neukončené volania funkcie *quicksort()*. V pravej časti obrázka (Obr. 3) sa ďalej nachádza aj vizualizácia delenia postupnosti triedených prvkov (*partitioning*) na základe náhodne zvoleného deliaceho prvku (*median*).

### B. Dynamické programovanie

Pre ilustráciu využitia metódy dynamického programovania pri návrhu efektívnych algoritmov je vytvorená vizualizácia algoritmu pre nájdenie minimálnej ceny násobenia postupnosti matíc (*Chain matrix multiplication*). Riešenie bez využitia dynamického programovania má exponenciálnu zložitosť a nie je preto vhodné pre praktické použitie pri vyššom počte matíc.

```

begin
  for i <- 1 until n do m[i][i] <- 0;
  for l <- 1 until n-1 do
    for i <- 1 until n-l do
      begin
        j <- i + l;
        for k <- i until k < j do
          begin
            q <- m[i][k]+m[k+1][j]+r[i-1]*r[k]*r[j];
            if(q < m[i][j])
              m[i][j] <- q;
          end
        end
      end
    end
  write m[l][n]
end
    
```

Počet matíc: n = 4    Rozmery matíc:

r[0]	r[1]	r[2]	r[3]	r[4]
10	20	50	1	100

Hodnoty premenných: l: 2 i: 1 j: 3 k: 2 q: 10500

m[1][1]=0	m[2][2]=0	m[3][3]=0	m[4][4]=0
m[1][2]=10000	m[2][3]=1000	m[3][4]=5000	

min(1200, 10500)

Obr. 4 Určenie minimálnej ceny násobenia postupnosti matíc.

Vizualizovaný algoritmus s využitím metódy dynamického programovania (Obr. 4) však má asymptotickú časovú zložitosť  $O(n^3)$  a je teda pre riešenie uvedeného problému prakticky použiteľný. Okrem vizualizácie pseudokódu algoritmu je zobrazený počet matíc, ich rozmery, aktuálne hodnoty premenných a postupne je vytváraná tabuľka minimálnych cien pre násobenie jednotlivých podpostupností matíc. Ak sa pri výpočte ceny násobenia určitej postupnosti matíc požaduje výpočet ceny násobenia niektorej jej podpostupnosti, tento výpočet sa algoritmom opakovane nevykonáva, ale použije sa skôr vypočítaná minimálna cena z tabuľky.

Výber algoritmu: MinMax

```

pair MinMax(A[], int L, int H) {
  pair ret1, ret2, ret;
  if(L == H)
  {
    ret.min = A[L];
    ret.max = A[L];
  }
  else if(L == H-1)
  {
    ret.min = MIN(A[L], A[H]);
    ret.max = MAX(A[L], A[H]);
  }
  else
  {
    ret1 = MinMax(A, L, (L+H)/2);
    ret2 = MinMax(A, (L+H)/2+1, H);
    ret.max = MAX(ret1.max, ret2.max);
    ret.min = MIN(ret1.min, ret2.min);
  }
  return ret;
}
    
```

0	1	2	3	4	5	6	7	8	9	(Index)
51	10	46	95	39	28	81	35	3	85	

MinMax L: 0 H: 9  
 MinMax L: 0 H: 4 Min: 10 Max: 95  
 MinMax L: 0 H: 2 Min: 10 Max: 51  
 MinMax L: 0 H: 1 Min: 10 Max: 51  
 MinMax L: 2 H: 2 Min: 46 Max: 46  
 MinMax L: 3 H: 4 Min: 39 Max: 95  
 MinMax L: 5 H: 9  
 MinMax L: 5 H: 7  
 MinMax L: 5 H: 6 Min: 28 Max: 81  
 MinMax L: 7 H: 7 Min: 35 Max: 35

Zásobník volaní: MinMax(A,0,9) MinMax(A,5,9) MinMax(A,5,7)

Informácie  
 Metóda: rekurzívna aplikácia procedúry MinMax s využitím techniky Rozdeľuj a panuj.  
 Procedúrou je pri každom volaní vrátená dvojica prvkov (minimálny a maximálny) z príslušného intervalu.  
 Zložitosť: T(n) = 3(n/2)-2.

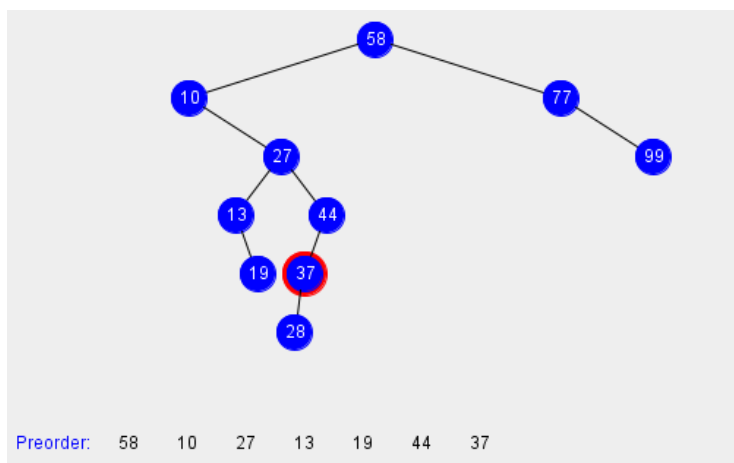
Obr. 5 Efektívne nájdenie minimálneho a maximálneho prvku.

### C. Rozdeľuj a panuj

Vizualizácia algoritmu pre efektívne nájdenie minimálneho a maximálneho prvku postupnosti (*MinMax*) je vytvorená pre ozrejmienie návrhu algoritmov s využitím metódy dekompozície (*Rozdeľuj a panuj*). Obrázok Obr. 5 v tomto prípade približuje celkové používateľské rozhranie aplikácie VizAlgo. Okrem samotnej vizualizácie algoritmu obsahuje aj pole s doplňujúcimi informáciami o danom algoritme a skupinu ovládacích prvkov. Prezentovaný algoritmus je rekurzívny a je pre neho dostupná aj vizualizácia zásobníka volaní funkcie *MinMax()*. Postupne sú zobrazované jednotlivé podpostupnosti (ohraničené hodnotami premenných L a H), na ktorých je algoritmom určované čiastkové minimum a maximum. Ak sú hodnoty minima a maxima pre danú podpostupnosť už známe, sú tieto hodnoty v rámci vizualizácie doplnené (zvýraznené červenou farbou). Ako posledné sú doplnené tieto hodnoty pre celú postupnosť (ohraničenú hodnotami L=0 a H=9 v prípade ako na Obr. 5).

### D. Rekurzívne algoritmy

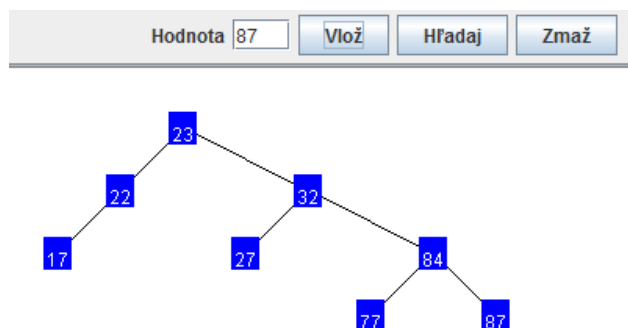
Okrem vyššie spomenutých rekurzívnych algoritmov (*Quicksort* a *MinMax*), ktoré boli zaradené v iných skupinách, je v rámci platformy dostupný aj rekurzívny prechod binárnym stromom stratégiou inorder, preorder a postorder. Na obrázku Obr. 6 je vizualizácia prechodu binárnym stromom stratégiou preorder. Aktuálne navštívený vrchol stromu je zvýraznený červenou farbou, pričom jeho hodnota je pridaná do zoznamu hodnôt už navštívených prvkov.



Obr. 6 Vizualizácia prechodu binárnym stromom stratégiou preorder.

### E. Vizualizácia údajových štruktúr

Z vizualizácií údajových štruktúr boli vytvorené vizualizácie abstraktných údajových typov (ADT) *Stack* (zásobník), *Queue* (front) a *BVS* (binárny vyhľadávací strom).



Obr. 7 Binárny vyhľadávací strom.

Pre ADT binárny vyhľadávací strom uvádzame ukážku z jeho vizualizácie (Obr. 7). Dostupné sú operácie pre vloženie, vyhľadanie a odobratie vloženého prvku. Táto vizualizácia prebieha v samostatnom okne, ktoré je otvorené po jej výbere v hlavnom okne programu.

#### IV. ZÁVER

V príspevku bol prezentovaný prehľad aktuálneho stavu vizualizačnej platformy VizAlgo. Aplikácia bola prakticky využitá ako podporný nástroj v rámci výučby predmetu Údajové štruktúry a algoritmy na FEI TU v Košiciach. Niektoré podnety pre využitie vizualizácie a ďalší vývoj aplikácie boli získané od poslucháčov uvedeného predmetu formou ankety [11]. Hoci počet vytvorených vizualizácií sa od času realizácie prieskumu formou ankety zvýšil, v rámci predmetu naďalej existujú tematické oblasti, ktoré nie sú pokryté platformou VizAlgo. Tieto oblasti sú vyučované tradičným spôsobom, prípadne sú využívané iné dostupné vizualizácie. Medzi perspektívne oblasti ďalšieho výskumu a vývoja tak patrí doplnenie zásuvných modulov s vizualizáciami ďalších algoritmov a údajových štruktúr, návrh a implementácia netradičných foriem vizualizácie, ako aj začlenenie podpory nových možností vizualizácie do jadra platformy VizAlgo.

#### LITERATÚRA

- [1] Computer Science Curriculum 2008, Association for Computing Machinery (ACM). Available: <http://www.acm.org/education/curricula>
- [2] Diehl, S.: Software visualization: Visualizing the Structure, Behaviour, and Evolution of Software. Springer, New York, 2007, 187 p.
- [3] Hundhausen, C. D., Douglas, S. A., and Stasko, J. T.: A Meta-Study of Algorithm Visualization Effectiveness, Journal of Visual Languages and Computing, Vol. 13, 2002.
- [4] JSPF: Java Simple Plugin Framework. Available: <http://code.google.com/p/jspf/>
- [5] Khuri, S.: Designing Effective Algorithm Visualizations, First International Program Visualization Workshop, ITiCSE, Finland, 2000. Available: <http://www.cs.sjsu.edu/~khuri/invited.html>
- [6] Kukfa, L.: Algorithm Operation Visualization, DCI FEEI TU of Košice, Bachelor thesis, 2014 (in Slovak).
- [7] Lazaridis, V., Samaras, N. and Sifaleras, A.: An empirical study on factors influencing the effectiveness of algorithm visualization. Computer Applications in Engineering Education, Vol. 21, 2013.
- [8] Mehlhorn, K., Sanders, P.: Algorithms and Data Structures, Springer-Verlag, 2008.
- [9] Šajko, A.: Algorithm Visualization, DCI FEEI TU of Košice, Bachelor thesis, 2012 (in Slovak).
- [10] Šimoňák, S.: Algorithm Visualization Using the VizAlgo Platform, Acta Electrotechnica et Informatica, Vol.13, No.2 (2013).
- [11] Šimoňák, S.: *Using Algorithm Visualizations in Computer Science Education*, Central European Journal of Computer Science (CEJCS), Versita, ISSN:1896-1533 (to be published).
- [12] Tudoreanu, M. E., Wu, R., Hamilton-Taylor, A., Kraemer, E.: Empirical Evidence that Algorithm Animation Promotes Understanding of Distributed Algorithms, IEEE 2002 Symposium on Human Centric Computing Languages and Environments, 2002.

# Verifying Secrecy and Authenticity of Security Protocols

*Martin TOMÁŠEK*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

`martin.tomasek@tuke.sk`

**Abstract** – We introduce security protocols by analyzing and verifying their properties. We use spi-calculus, an extension of the  $\pi$ -calculus, that enables us to consider cryptographic issues in more details. In this work we represent the security protocol as a process and we use the behavioral equivalences for describing secrecy and authenticity properties. Our goal is to design the practical procedure for verification of security protocols.

**Keywords** – protocol, spi-calculus, secrecy, authenticity,  $\pi$ -calculus.

## I. INTRODUCTION

Cryptographic protocols are used today to provide security in various applications. Cryptographic protocols are rules for exchange of messages between participants, and rely on cryptographic algorithms like encryption and decryption. Experience has shown that even very simple protocols which seem secure may have subtle flaws, even if the underlying cryptographic algorithms are secure. An extension of the  $\pi$ -calculus, the spi-calculus [1], was proposed as a formal notation for describing and reasoning about cryptographic protocols.

The objective of our work is to find a practical method of modeling and verifying cryptographic protocols using spi-calculus and validate it on specific communicating protocols. We analyze cryptographic protocols and their security properties. By means of basic knowledge about process algebras we use spicalculus for specification of cryptographic protocols. We develop and evaluate common formal method for the verification of cryptographic protocols.

## II. EXPRESSING SECURITY PROTOCOLS

A protocol  $P = C^* \cup C$ , where clauses in  $C$  use symbols from  $\Sigma$ , predicates from  $P^* \cup P$ , and contain predicates from  $P^*$  only in the body. We can write  $\bar{c}(M).P$  to denote a process that sends the message  $M$  on channel  $c$  after which it executes the process  $P$ . Then  $c(x)M$  denotes a process that is listening on the channel  $c$  and if it receives some message  $M$  on this channel then it will execute the process  $Q[M/x]$ . We may compose these two processes in parallel to get a bigger process, denoted as  $\bar{c}(M).P \mid c(x).Q$ . Now the two smaller processes may communicate on the channel  $c$  after which they will execute the process  $P \mid Q[M/x]$  [2]. The cryptographic protocol is communicating protocol, which uses the cryptography to achieve security goals. Basic cryptographic algorithms are DES, RSA, and DSA, and may be vulnerable if key is too short.

### A. Abstract Syntax of the Calculus

The abstract syntax of the spi-calculus [1] is divided into two parts, terms and processes.

$L, M, N ::=$	<b>Terms</b>
$n$	Name
$(M, N)$	Pair
$0$	Zero
$suc(M)$	Successor
$x$	Variable
$\{M\}_N$	Shared key encryption
$P, Q, R ::=$	<b>Processes</b>
$M(N).P$	Output - process is ready to output on channel $m$
$M(x).P$	Input - process is ready to input from channel $m$
$P \mid Q$	Composition - behaves as process $P$ and $Q$ running in parallel
$(vn)P$	Restriction is a process that makes a new, private name $n$ , which may occur in $P$
$!P$	Replication behaves as a finite number of copies of $P$ running in parallel
$[M \text{ is } N]P$	Match behaves as $P$ if the terms $M$ and $N$ are the same; otherwise it is stuck (it does nothing)
$0$	Nil process does nothing
$let (x, y) = M \text{ in } P$	Pair splitting processes $let (x, y) = M \text{ in } P$ behaves as $P[N/x][L/y]$ if the term $M$ is the pair $(N, L)$
$case M \text{ of } 0 : P \ suc(x) : Q$	Integer case behaves as $P$ if the term $M$ is 0, as $Q[N/x]$ if $M$ is $suc(N)$
$case L \text{ of } \{x\}_N \text{ in } P$	Shared key decryption

### B. Semantic of the Calculus

Let  $fn(M)$  and  $fn(P)$  be a set of free names in term  $M$  and process  $P$ . Let  $fv(M)$  and  $fv(P)$  be the set of free variables in term  $M$  and process  $P$ . Closed processes are processes without any free variables. [3]

**Reaction relation;**  $P \rightarrow Q$  means that there exists a reaction between subprocesses of  $P$  such that the whole can step to process  $Q$ :

$$\bar{m}(N).P \mid m(x).Q \rightarrow P \mid Q[N/x] \quad \text{Interaction}$$

Then we define the **reduction relation**  $>$  on closed processes:

$$\begin{aligned} !P &> P \mid !P && \text{Replication} \\ [M \text{ is } M]P &> P && \text{Match} \\ let (x, y) = (M, N) \text{ is } P &> P[M/x][N/y] && \text{Let} \\ case 0 \text{ of } 0 : P \ suc(x) : Q &> P && \text{Zero} \\ case suc(M) \text{ of } 0 : P \ suc(x) : Q &> Q[M/x] && \text{Successor} \\ case \{M\}_N \text{ of } \{x\}_N \text{ in } P &> P[M/x] && \text{Decrypt} \end{aligned}$$

**Structural equivalence** is a relation on closed processes that satisfies the following rules and equation:

$$\begin{array}{ll} P \mid 0 \equiv P & \text{Nil} \\ P \mid Q \equiv Q \mid P & \text{Commitment} \\ P \mid (Q \mid R) \equiv (P \mid Q) \mid R & \text{Association} \\ (vm)(vn)P \equiv (vn)(vm)P & \text{Switch} \\ (vn)0 \equiv 0 & \text{Drop} \\ (vn)(P \mid Q) \equiv P \mid (vn)Q \text{ if } n \notin fn(P) & \text{Extrusion} \end{array}$$

$$\begin{array}{lll} \text{Reduction} & \text{Reflection} & \text{Symmetry} \\ \frac{P > Q}{P \equiv Q} & \frac{}{P \equiv P} & \frac{P \equiv Q}{Q \equiv P} \\ \text{Transitivity} & \text{Parametrization} & \text{Restriction} \\ \frac{P \equiv Q \ Q \equiv R}{P \equiv R} & \frac{P \equiv P'}{P \mid Q \equiv P' \mid Q} & \frac{P \equiv P'}{(vm)P \equiv (vm)P'} \end{array}$$

With these rules we can complete reaction rules as follow:

$$\frac{P \equiv P' \ P' \rightarrow Q' \ Q' \equiv Q}{P \rightarrow Q} \quad \frac{P \rightarrow P'}{P \mid Q \rightarrow P' \mid Q} \quad \frac{P \rightarrow P'}{(vn)P \rightarrow (vn)P'}$$

Abadi and Gordon [1] use **testing equivalence** as the notion of equivalence. Two processes are testing equivalent, written  $P \simeq Q$ , if they are indistinguishable to any other process. For

specification of testing equivalence [4] we first define barbs. **Barbs** define a predicate describing the channels, where output process can communicate. A barb  $\beta$  is an input or an output channel, where output channels are marked by a barb  $\bar{m}$ .  $P$  exhibits barb  $\beta$ , written  $P \downarrow \beta$ , is defined:

$$\begin{array}{l}
 m(x).P \downarrow m \quad \text{Input} \\
 \bar{m}\langle M \rangle.P \downarrow \bar{m} \quad \text{Output} \\
 \\
 \begin{array}{ccc}
 \text{Barb Parametrization} & \text{Barb Restriction} & \text{Barb Structural} \\
 \frac{P \downarrow \beta}{P \mid Q \downarrow \beta} & \frac{P \downarrow \beta \ \beta \notin \{m, \bar{m}\}}{(vm)P \downarrow \beta} & \frac{P \equiv Q \ Q \downarrow \beta}{P \downarrow \beta}
 \end{array}
 \end{array}$$

Test is a closed process  $R$  and a barb  $\beta$ . The process  $R$  is trying to see if the tested process can be made to exhibit barb  $\beta$ :

$$P \sqsubseteq Q = \text{for any test } (R, \beta), \text{ if } (P \mid R) \downarrow \beta \text{ then } (Q \mid R) \downarrow \beta \quad \text{Testing Preorder}$$

$$P \simeq Q = P \sqsubseteq Q \text{ and } Q \sqsubseteq P \quad \text{Testing Equivalence}$$

The idea about testing equivalence builds De Nicola and Hennesy [5].

### III. SECURITY PROPERTIES AND VERIFICATION PROCEDURE

For the verification of cryptographic protocols it is useful first define security properties [6] of these protocols.

**Secrecy:**  $M$  is secret if a session that contains  $M$  is indistinguishable from any session containing some data  $M_0$  in place of  $M$  (observational equivalence property). Global secrecy is when a message is secret all the time. Local secrecy is when a message is secret till the corresponding session has not ended.

**Authenticity:** If  $A$  accepts a message  $M$  as coming from  $B$  then  $B$  actually sent  $M$ . If  $A$  received a message of form  $M_1$  then  $B$  sent a message of form  $M_2$ . If  $A$  got a message of form  $M$  then  $B$  was active. If  $A$  has got a message  $M$   $n$  times then  $B$  sent it  $n$  times.

In this project we want to proceed verification of cryptographic protocols by means of validation of the secrecy and the authenticity. We define the safety property.

**Definition 4.1:** Safety

- **Authenticity:**  $B$  always replies  $F$  to the message  $M$  that  $A$  sends; an attacker cant cause  $B$  to apply  $F$  to some other message.
- **Secrecy:** The message  $M$  can't be read in transit from  $A$  to  $B$ ; if  $F$  doesn't reveal  $M$ , then the whole protocol doesn't reveal  $M$ .

Protocol is safe only if both conditions, authenticity and secrecy, are satisfied. In summary, we have:

$$\begin{array}{ll}
 Inst(M) \simeq Inst_{spec}(M), \text{ for all } M & \text{Authenticity} \\
 Inst(M) \simeq Inst(M') \text{ if } F(M) \simeq F(M'), \text{ for all } M, M' & \text{Secrecy}
 \end{array}$$

#### A. Verification Procedure

We designed following procedure to verify the safety properties of the communication protocols:

- 1) Write the protocol into convenient form. The best is writing it with messages.
- 2) Make the spi-calculus description of this protocol.
- 3) Make specification from description of this protocol.
- 4) Verify authenticity:
  - Make specification for authenticity.
  - Verify authenticity by exhibiting auxiliary equivalences (strong bisimilarity, barbed equivalence, and barbed congruence).
- 5) Verify secrecy:
  - Prove restricted version of secrecy property  
 $Inst(M) \simeq Inst(M') \text{ if } F(x) \text{ is } \bar{c}\langle * \rangle.$
  - Prove full secrecy property  $Inst(M) \simeq Inst(M') \text{ if } F(M) \simeq F(M')$  using auxiliary equivalences.
- 6) If both authenticity and secrecy are valid, then the protocol is secure.

### IV. EXAMPLES

Two principals  $A$  and  $B$  share the key  $K_{AB}$ , we assume there is a public channel  $c_{AB}$  that  $A$  and  $B$  use for communication. The protocol is simply that  $A$  sends a message  $M$  under  $K_{AB}$  to  $B$ , on  $c_{AB}$ .

To verify the safety properties of the protocol we use the above procedure.

- 1) Message 1  $A \rightarrow B : \{M\}_{K_{AB}} \text{ on } c_{AB}$

2) Specification in spi-calculus

$$\begin{aligned} A(M) &= \overline{c_{AB}}\langle \{M\}_{K_{AB}} \rangle \\ B_{spec} &= c_{AB}(x).case\ x\ of\ \{y\}_{K_{AB}}\ in\ F(y) \\ Inst_{spec}(M) &= (vc_{K_{AB}})(A(M) \mid B_{spec}(M)) \end{aligned}$$

3) The main definitions are:

$$\begin{aligned} Inst(M) &= (vc_{AB})(\overline{c_{AB}}\langle M \rangle.0 \mid c_{AB}(x).F(x)) \\ Inst_{spec}(M) &= (vc_{AB})(\overline{c_{AB}}\langle M \rangle.0 \mid c_{AB}(x).F(M)) \end{aligned}$$

4) **Proposition 5.1:** For any closed term  $M$ ,  $Inst(M) \simeq Inst_{spec}(M)$ .

Only commitments of  $Inst(M)$  and  $Inst_{spec}(M)$  are:

$$\begin{aligned} Inst(M) &\xrightarrow{\tau} (vc_{AB})(0 \mid F(M)) \\ Inst_{spec}(M) &\xrightarrow{\tau} (vc_{AB})(0 \mid F(M)) \end{aligned}$$

From definition of barbed congruence we know, that strong bisimilarity implies barbed congruence and barbed congruence implies testing equivalence.

$$\begin{aligned} Inst(M) &\sim_s Inst_{spec}(M) \\ Inst(M) &\sim Inst_{spec}(M) \\ Inst(M) &\simeq Inst_{spec}(M) \end{aligned}$$

5) First we prove restricted version of secrecy property.

**Lemma 5.2:**  $Inst(M) \simeq Inst(M')$  if  $F(M')$  is  $\bar{c}(\ast)$ , for any closed terms  $M$  and  $M'$ . Only commitment of  $Inst(N)$  is:  $Inst(N) \xrightarrow{\tau} (vc_{AB})(0 \mid \bar{c}(\ast))$  and so clearly  $Inst(M) \sim_s Inst(M')$ . Like in previous,  $Inst(M) \simeq Inst(M')$ .

Now we can make calculation of full secrecy property,  $Inst(M) \simeq Inst(M')$  if  $F(M) \simeq F(M')$ . In special case, where  $F(x)$  is  $\bar{c}(\ast)$  we can write  $Inst(M, (x)\bar{c}(\ast))$ . We assume that  $c$  is a fresh name and  $y$  fresh variable and we write  $\tau.F(N)$  for  $(vc)(\bar{c}(\ast) \mid c(y).F(N))$ . Only commitments are:

$$\begin{aligned} (vc)(c_{AB}(x).\bar{c}(\ast) \mid c(y).F(N)) &\xrightarrow{c_{AB}} (x)\tau.F(N) \\ c_{AB}(x).\tau.F(N) &\xrightarrow{c_{AB}} (x)\tau.F(N) \end{aligned}$$

From these, we have:  $(vc)(c_{AB}(x).\bar{c}(\ast) \mid c(y).F(N)) \sim_s c_{AB}(x)\tau.F(N)$ .

As follows using proposition  $F(N) \simeq \tau.F(N)$ , facts that testing equivalence is congruence and that strong bisimilarity implies testing equivalence, we have:

$$\begin{aligned} Inst_{spec}(N) &= (vc_{AB})(\overline{c_{AB}}\langle N \rangle.0 \mid c_{AB}(x).F(N)) \\ &\simeq (vc_{AB})(\overline{c_{AB}}\langle N \rangle.0 \mid c_{AB}(x).(\tau.F(N))) \\ &\simeq (vc_{AB})(\overline{c_{AB}}\langle N \rangle.0 \mid (vc)(c_{AB}(x).\bar{c}(\ast) \mid c(y)F(N))) \\ &\equiv (vc)((vc_{AB})(\overline{c_{AB}}\langle N \rangle.0 \mid (vc)(c_{AB}(x).\bar{c}(\ast) \mid c(y)F(N))) \\ &= (vc)(Inst(N, (x)\bar{c}(\ast)) \mid c(y).F(N)) \end{aligned}$$

And we obtain equation:  $Inst_{spec} \simeq (vc)(Inst(N, (x)\bar{c}(\ast)) \mid c(y).F(N))$ . With this equation, Lemma 5.2, Proposition 5.1 and assumption  $F(M) \simeq F(M')$  we can make following calculation.

$$\begin{aligned} Inst(M) &\simeq Inst_{spec}(M) \\ &\simeq (vc)(Inst(M, (x)\bar{c}(\ast)) \mid c(y).F(M)) \\ &\simeq (vc)(Inst(M', (x)\bar{c}(\ast)) \mid c(y).F(M')) \\ &\simeq Inst_{spec}(M') \\ &\simeq Inst(M') \end{aligned}$$

6) Authenticity and secrecy property are valid, protocol is secure.

## V. CONCLUSION

This work describes verification of cryptography protocols with emphasis on authenticity and secrecy properties using spi-calculus. The main task was to design a common procedure of the verification, which can be applied on any cryptographic protocol. Presented results are based on Abadi's and Gordon's testing equivalence and auxiliary equivalences [3]. This approach is more suitable for automation than solution designed by Woo and Lam [2]. Future extension of this work may be a software implementation of designed procedure.

## ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10.

#### REFERENCES

- [1] M. Abadi and A. D. Gordon, "A calculus for cryptographic protocols: The spi calculus," *Information and Computation*, vol. 148, no. 1, pp. 1 – 70, 1999. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0890540198927407>
- [2] T. Y. C. Woo and S. S. Lam, "A semantic model for authentication protocols," 1993.
- [3] M. Abadi and A. D. Gordon, "Reasoning about cryptographic protocols in the spi calculus," in *In CONCUR'97: Concurrency Theory*. Springer-Verlag, 1997, pp. 59–73.
- [4] R. Cleaveland and M. Hennessy, "Testing equivalence as a bisimulation equivalence," 1993.
- [5] R. de Nicola and M. Hennessy, "Testing equivalences for processes," *Theoretical Computer Science*, vol. 34, pp. 83–133, 1984.
- [6] A. D. Gordon and A. Jeffrey, "Authenticity by typing for security protocols," *Journal of Computer Security*, p. 2003.



# OrCAD PSpice – model výkonového MOSFET tranzistora

<sup>1</sup>Milan Lacko, <sup>2</sup>Peter Girovský

<sup>1</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovensko

<sup>2</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovensko

<sup>1</sup>milan.lacko@tuke.sk, <sup>2</sup>peter.girovsky@tuke.sk

**Abstrakt** — Simulácia výkonových obvodov meničov predstavuje v dnešnej dobe nevyhnutnú súčasť pri navrhovaní meničov. Na takúto simuláciu je výhodné použiť program OrCAD PSpice, ktorý je vhodný pre simulovanie správania sa celého meniča, ale aj jednotlivých výkonových súčiastok. Tento článok sa podrobne venuje vytváraniu vlastného simulačného modelu výkonového MOSFET tranzistora v prostredí OrCAD PSpice a opisu parametrov pri vytváraní modelu tranzistora.

**Keywords** — model, MOSFET, OrCAD PSpice, simulácia

## I. ÚVOD

Pri návrhu a dimenzovaní meničov a výkonových obvodov je nevyhnutné, aby sme správanie sa navrhovaného obvodu či celého meniča najprv odsimulovali, a tak si objasnili, ako daný obvod pracuje. Samozrejme, zaujíma nás aj správanie sa jednotlivých súčiastok. Prúdy a napätia v jednotlivých stavoch, vznik možných prepätí či celkové straty pri spínaní výkonových súčiastok. Preto je nevyhnutné v simulácii používať čo najreálnejšie modely jednotlivých súčiastok, pre ktoré sme menič nadimenzovali, a ktoré neskôr budú použité pri realizácii meniča. V tomto článku sa budeme zaoberať simuláciou v prostredí OrCAD PSpice, konkrétne simuláciou a vytváraním vlastného modelu výkonového tranzistora MOSFET. Nami zvolený tranzistor je IRF3205.

## II. VYTVÁRANIE VLASTNÉHO SIMULAČNÉHO MODELU TRANZISTORA MOSFET

Nakoľko vývoj polovodičových súčiastok napreduje, aj konštruktéri chcú používať pre svoje meniče tieto lepšie súčiastky. Nato je však potrebné, aby spolu s novými súčiastkami prichádzali aj ich simulačné modely, aby bolo možné detailne simulovať správanie sa meniča s týmito novými súčiastkami. Niektorí výrobcovia pri uvedení novej polovodičovej súčiastky na trh hneď uvádzajú aj hotový simulačný model tejto súčiastky vhodný napr. práve pre spomínaný OrCAD PSpice. Avšak stále je veľké množstvo nových súčiastok, ku ktorým nie je výrobcom vytvorený model, prípadne je tento model veľmi zjednodušený a odzrkadľuje len tie najzákladnejšie parametre súčiastky. Vtedy nie je možné pri simulácii použiť nami zvolenú polovodičovú súčiastku, nie je možné dostatočne presne zistiť napr. straty súčiastky, správanie sa pri spínaní vyššími frekvenciami či prípadné oscilačné deje, a je potrebné vytvoriť si vlastný simulačný model.

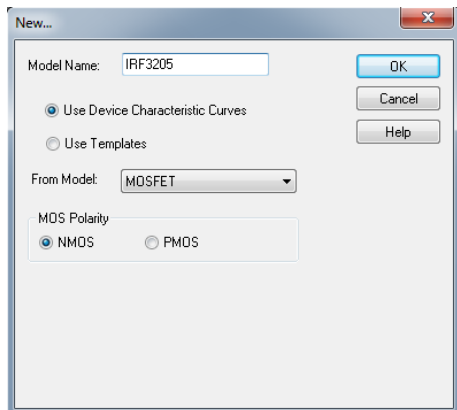
### A. Vytvorenie simulačného modelu tranzistora MOSFET podľa katalógového listu

Ak máme k dispozícii od výrobcu katalógový list konkrétneho tranzistora, ktorého model chceme vytvoriť, je potrebné vychádzať z tohto dokumentu. Je potrebné zadať nie len statické parametre MOSFET tranzistora, ale aj dynamické parametre, ktoré sú väčšinou uvedené vo forme grafu alebo charakteristiky.

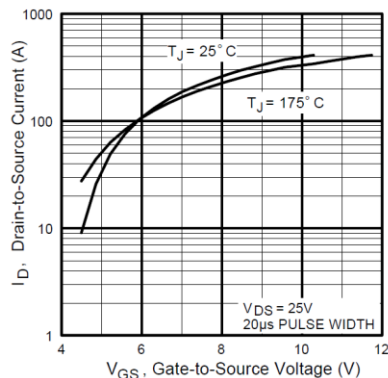
V našom prípade budeme vytvárať model tranzistora IRF3205. Je to MOSFET tranzistor a jeho základné parametre sú:  $U_{DS} = 55$  V,  $I_D = 110$  A,  $R_{DS(on)} = 8$  m $\Omega$ ,  $t_{on} = 115$  ns,  $t_{off} = 115$  ns.

Model budeme vytvárať v simulačnom programe Model Editor, ktorý je súčasťou softvéru

OrCAD PSpice, určeného na simuláciu elektrotechnických obvodov. Po spustení programu si vytvoríme nový projekt, kliknutím na *File/ New*. Pre vytvorenie nového modelu si musíme vybrať model súčiastky, ktorý chceme vytvárať. V hlavnom menu klikneme na *Model/ New*, kde zadávame názov modelu, a či ho chceme vytvárať podľa charakteristík alebo pomocou šablóny. V našom prípade vyberiem vytváranie podľa charakteristík a následne v položke *From model* vyberieme model, aký chceme vytvoriť. V našom prípade je to model MOSFET tranzistora s n-kanálom. Úvodné okno model editora je na Obr. 1.

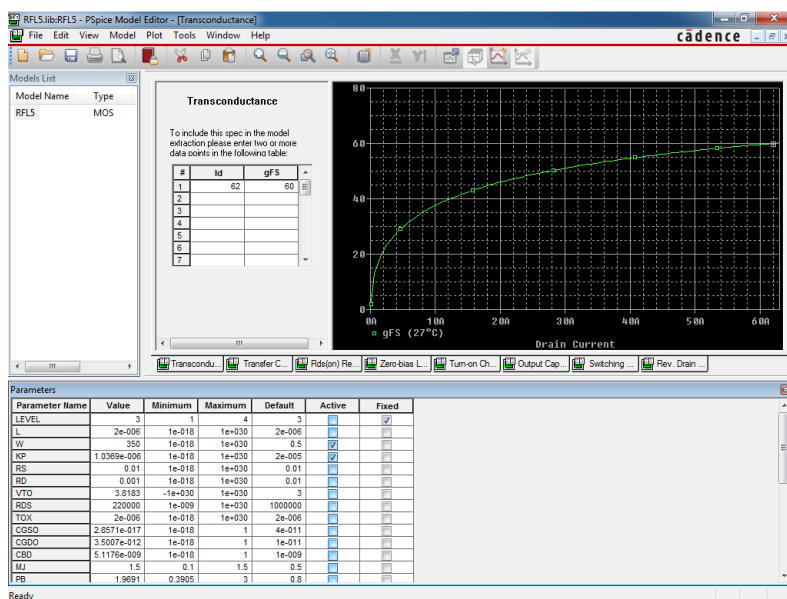


Obr. 1 Model Editor – úvodné okno



Obr. 2 Prenosová charakteristika tranzistora IRF3205

Po potvrdení kliknutím na tlačidlo OK, sa spustí modelovací editor, v ktorom môžeme začať modelovať MOSFET tranzistor. Po spustení modelovacieho editora vidíme tri okná. V prvom okne vľavo je Model list, kde vidíme názov a typ tranzistora.



Obr.3 Okno Model Editora – záložka Transconductance

V druhom, hlavnom, okne je zobrazená charakteristika strmosti s názvom *Transconductance*, ktorá sa prepína medzi ostatnými záložkami, do ktorých sa vkladajú charakteristiky a parametre, ktoré sú v katalógovom liste. V dolnej časti sa nachádza okno s modelovacími parametrami, ktoré sa budú počas vytvárania modelu meniť.

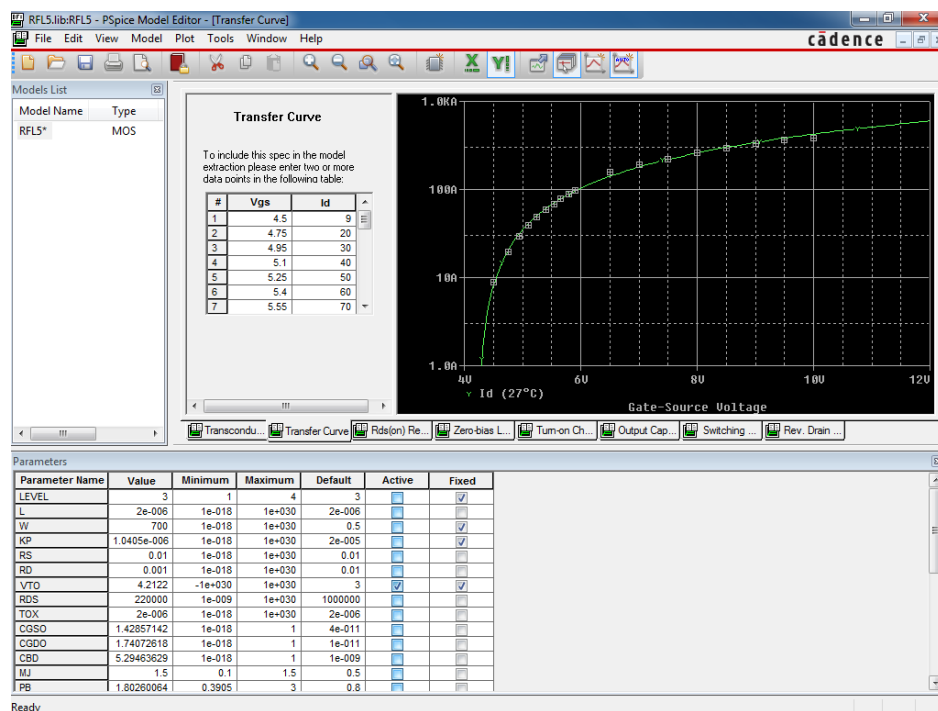
V katalógovom liste MOSFET tranzistora si nájdeme strmost' (*Forward Transconductance*)  $g_{fs}$ , ktorej hodnota je 44. Pri výkonových tranzistoroch je zadaná len jedna hodnota. Pre tento prípad bola zadaná iba minimálna hodnota, preto bolo vhodné ju navýšiť na hodnotu 60. V katalógovom liste sa nachádzajú hodnoty jednotlivých parametrov, ktoré sú zadané ako hraničné hodnoty (minimálne, maximálne) a typizované hodnoty. Hodnotu zapíšeme do príslušných položiek v Model Editore (viď. Obr. 3) a na nástrojovej lište si nájdeme *Extract*, na ktorý klikneme a vidíme, že sa nám zmenili hodnoty modelových parametrov W a KP, ktoré sú zaznačené v stĺpci *Active*. Následne parametre týchto zmenených hodnôt zaškrtneme na *Fixed*.

Znovu stlačíme tlačidlo *Extract*, a hodnoty týchto parametrov zaokrúhli. Hodnoty uvedených parametrov sa už počas zadávania ďalších hodnôt parametrov tranzistora MOSFET nebudú meniť.

Následne sa preklikneme na záložku *Transfer Curve*. V katalógovom liste MOSFET tranzistora si nájdeme prenosovú charakteristiku (Obr. 2). Z charakteristiky odčítame dostatočný počet bodov pri teplote 25°C a následne ich zapíšeme do príslušných políček v Model Editore (Obr. 4). Čím viac bodov odčítame, tým je charakteristika presnejšia, a tým sú presnejšie vypočítané parametre pre model tranzistora.

Vedľa tabuľky zapísaných hodnôt je zobrazený priebeh, ktorý sa vytvoril na základe zadaných hodnôt. Graf si upravíme prepnutím na logaritmickú mierku tak, že na nástrojovej lište klikneme na *Plot/Axis Settings...* Po zapísaní hodnôt sa krivka nezhodovala so zadávanými bodmi, a to preto, že v predchádzajúcej položke bol zadaný parameter strmosti ako minimálna hodnota, a preto je potrebné zvýšiť hodnotu. Vidíme, že po zvýšení hodnôt sa nám aj krivka prenosovej charakteristiky začala zhodovať s bodmi. Následne klikneme na *Extract*, zaškrtneme aktívny parameter VTO a znovu stlačíme *Extract*.

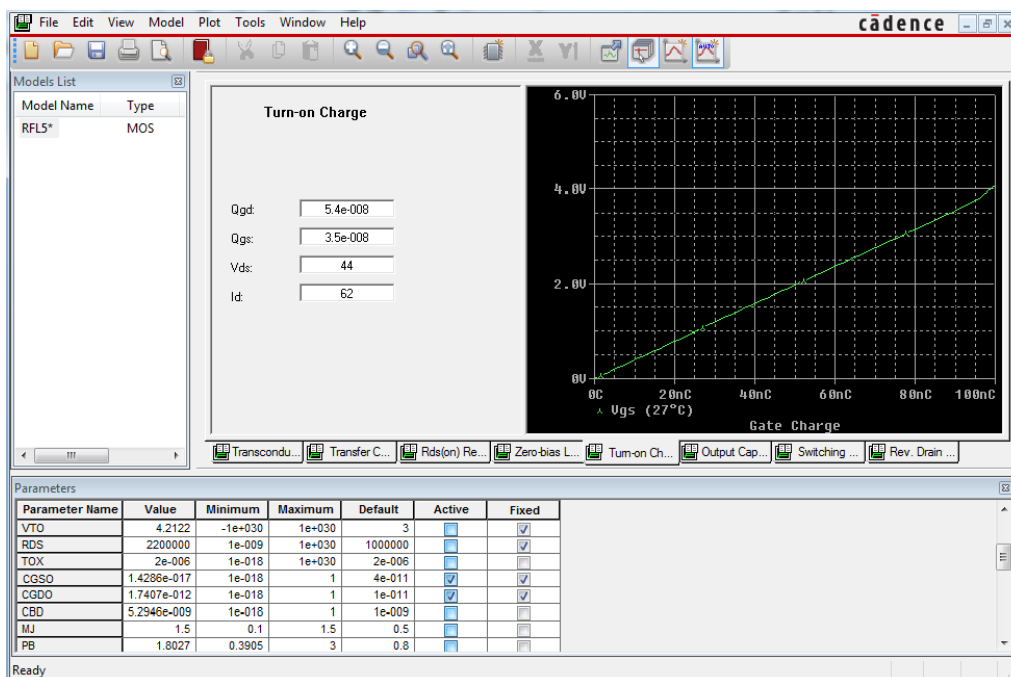
Klikneme na záložku *Rds(on) resistance*, kde doplníme požadované údaje podľa katalógového listu. Nájdeme si  $R_{DS(on)}$  – odpor tranzistora MOSFET v zapnutom stave, pri ktorom sú zadané i zvyšné hodnoty  $I_D$  a  $V_{GS}$ , ktoré sa nachádzajú v poslednom stĺpci tabuľky pod názvom *Podmienky (Conditions)*. Po zadaní hodnôt, v tomto prípade  $I_D = 62 \text{ A}$ ,  $R_{DS(on)} = 8 \text{ m}\Omega$ ,  $V_{GS} = 10 \text{ V}$ , klikneme na *Extract*, vidíme že sa nám zmenila hodnota aktívneho parametra *RD* a následne zaškrtneme aktívny parameter na *Fixed* a stlačíme *Extract*.



Obr. 4 Okno Model Editora – záložka Transfer Curve

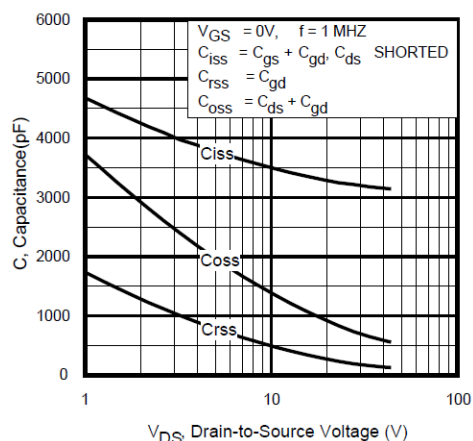
V záložke *Zero-bias Leakage* – zvodové nulové prepätie – zadáme hodnotu  $I_{DSS}$  a  $V_{DS}$ . V katalógovom liste si nájdeme hodnotu zvodového prúdu medzi kolektorom a emitorom (Drain to Source Leakage Current).  $I_{DSS} = 25 \mu\text{A}$  a napätia  $V_{DS} = 55\text{V}$  pri teplote 25°C. Zapíšeme ich a stlačíme *Extract*, zmení sa nám parameter *RDS*, zaškrtneme ho v poličku *Fixed* a stlačíme *Extract*.

Preklikneme sa na záložku *Turn-on Charge* – zapínací náboj (Obr. 5), kde budeme zadávať hodnoty parametrov z katalógu, ktoré nájdeme pod tým istým označením ako v Model Editore. Do položky *Qgd* zapíšeme náboj medzi hradlom a kolektorom, *Qgs* zapíšeme náboj medzi hradlom a emitorom a prislúchajúci prúd  $I_D$  a napätie  $V_{DS}$ , ktoré sa nachádza v podmienkach pre dané parametre. Pre použitý typ MOSFET tranzistora sú to hodnoty  $Q_{gs} = 35 \text{ nC}$ ,  $Q_{gd} = 54 \text{ nC}$ ,  $I_D = 62 \text{ A}$ ,  $V_{DS} = 44 \text{ V}$ . Po zadaní hodnôt stlačíme *Extract* a zafixujeme hodnoty.

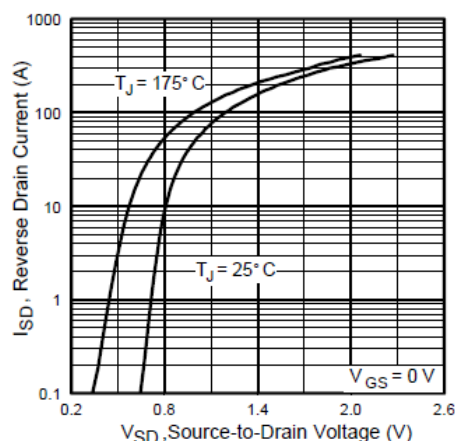


Obr.5 Okno Model Editora – záložka Turn-on Charge

Ďalej sa preklikneme na záložku Output Capacitance – výstupná kapacita – kde si nájdeme v katalogovom liste tranzistora charakteristiky závislosti kapacit od napätia  $V_{DS}$  (Obr. 6).



Obr. 6 Priebeh kapacity v závislosti od napätia  $V_{DS}$

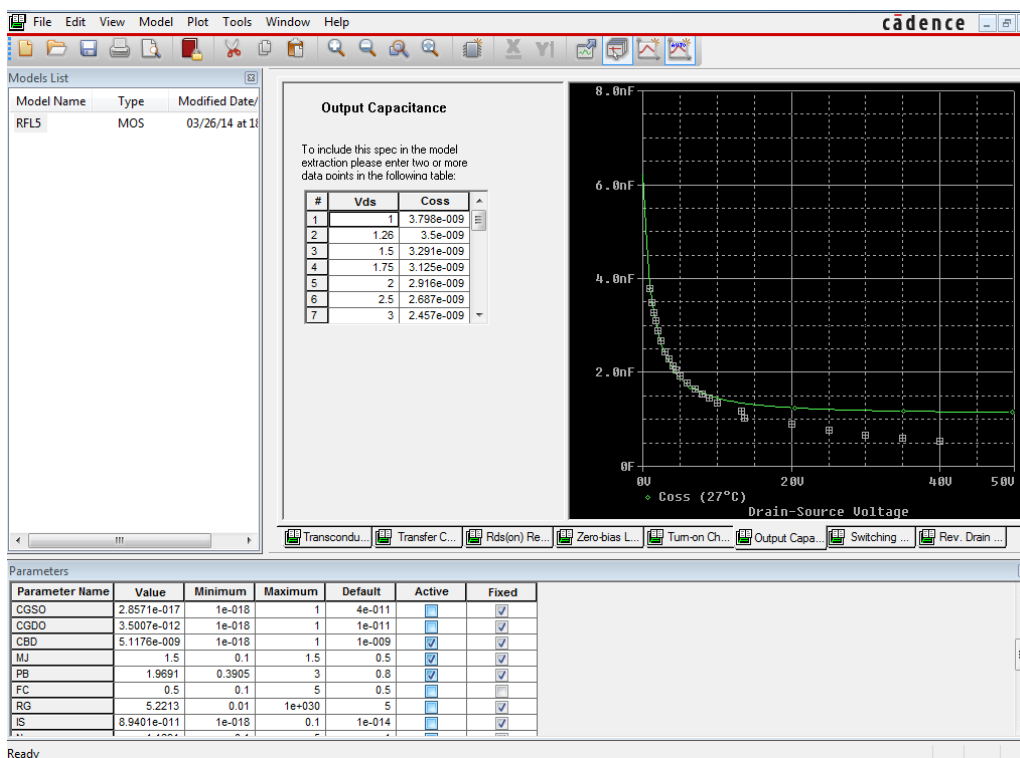


Obr. 7 Priebeh spätného prúdu  $I_{SD}$

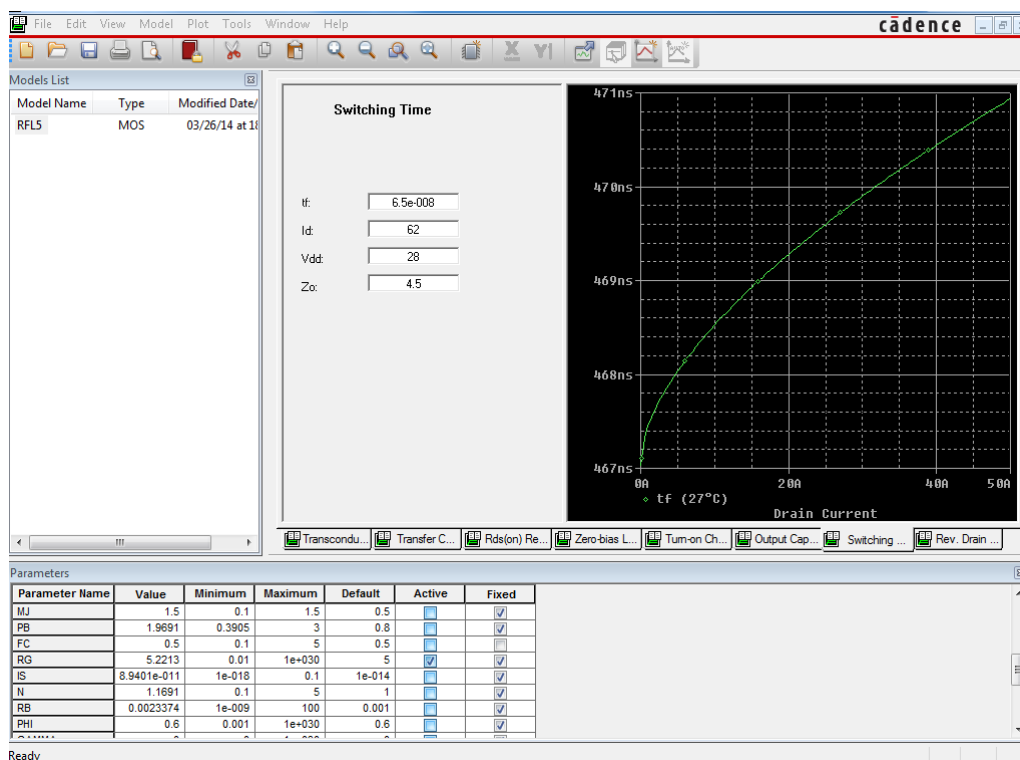
Vidíme, že v grafe sú znázornené viaceré kapacity, my budeme pracovať s výstupnou kapacitou  $C_{oss}$ . Priebeh si rozdelíme na potrebný počet bodov, ktoré zapíšeme do tabuľky v Model Editore. Znovu stlačíme *Extract*, zafixujeme aktívne parametre CBD, MJ, PB a ešte raz stlačíme *Extract* (Obr. 8).

Prekliknutím sa na záložku Switching Time (spínacie časy) sa nám zobrazí okno Model Editora podľa Obr. 9. V katalogovom liste si nájdeme dynamické parametre zvoleného tranzistora a z nich si vyberieme potrebný parameter  $t_f$  – dobu poklesu. Nájdeme si ho pod označením  $t_f$  (fall time), pri ktorom sa hneď nachádzajú aj ďalšie potrebné parametre v stĺpci s podmienkami, kde  $I_d = I_D, V_{dd} = V_{DD}$  a  $Z_o = R_G$ . Po dosadení parametrov postupujeme rovnako ako v predchádzajúcich prípadoch. Aktívnym parametrom je  $R_G$  parameter.

Nakoniec sa preklikneme na záložku Rev. Drain Current, kde budeme postupovať podobne ako pri prenosovej charakteristike a výstupnej kapacite. V katalogovom liste si nájdeme závislosť spätného kolektorového prúdu (Reverse Drain Current)  $I_{SD}$  od napätia medzi emitorom a kolektorom  $V_{SD}$  (Obr. 7). Odčítané hodnoty pri teplote  $25^\circ\text{C}$ , zapíšeme ich do tabuľky, stlačíme *Extract*, zafixujeme aktívne parametre IS, N, RB a znovu stlačíme *Extract* (Obr. 10).

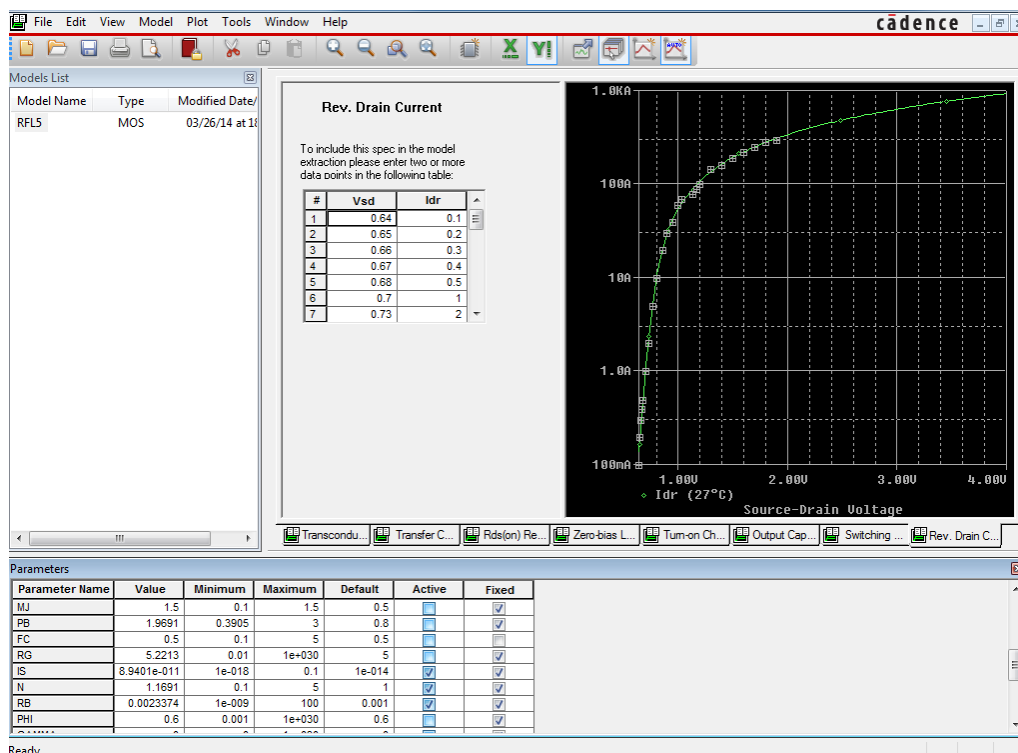


Obr. 8 Okno Model Editora – záložka Output Capacitance



Obr. 9 Okno Model Editora – záložka Switching Time

Týmto postupom sme vytvorili vlastný simulačný model konkrétneho MOSFET tranzistora. Aby sme mohli vytvorený model využívať v simulačnom programe PSpice je potrebné ho uložiť do knižnice PSpice. Môžeme to spraviť dvoma spôsobmi. Prvý spôsob je, že klikneme na *File/Save As* alebo tak, že klikneme priamo na ikonu *Save Library*. Nasledovne sa otvorí malé okno, kde zadáme umiestnenia a názov knižnice vytvoreného modelu. Ak chceme využívať vytvorený



Obr. 10 Okno Model Editora – záložka Reverse Drain Current

model MOSFET tranzistora pri simulácii je potrebné pre PSpice knižnicu vytvoriť súbor s príponou .olb. Súbor s príponou .olb vytvoríme v hlavnom menu kliknutím na *File/Model Import Wizard Capture*. Otvorí sa malé okno, v ktorom si vyberieme v položke *Enter Inputs Model Library* náš vytvorený model tranzistora a v položke *Enter Destination Symbol Library* miesto, kde sa uloží súbor s príponou .olb. Potvrdíme stlačením *OK* a nami vytvorený model tranzistora je pripravený na používanie v simuláciách.

## I. ZÁVER

V tomto článku je opísaný návod na vytvorenie vlastného simulačného modelu MOSFET tranzistora. Tento podrobný návod na vytvorenie vlastného modelu tranzistora v programe OrCAD PSpice nám dáva možnosť s vysokou presnosťou simulovať správanie sa meniča, kde sú použité tieto modely. Takáto simulácia dáva veľmi dobrý reálny pohľad na namáhanie súčiastok meniča, presnejšie vyčíslí straty jednotlivých tranzistorov a hlavne reálnejšie zobrazí dynamické správanie sa výkonových polovodičových meničov s ohľadom na vysoké frekvencie spínania. S kvalitným simulačným modelom môžeme dimenzovať kvalitné zariadenie bez zbytočného predimenzovania, a tak dokážeme ušetriť finančné prostriedky a čas vyčlenený na vývoj zariadenia.

## POĎAKOVANIE

Táto práca bola podporená Kultúrnou a edukačnou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky pod projektom č.: 042TUKE-4/2012 „Inovácia výučby riadenia mechatronických systémov“.

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10

## REFERENCIE

- [1] OrCAD PSpice A/D, Reference manual, dostupné na internete, <http://www.electronics-lab.com/downloads/schematic/013/tutorial/PSPICREF.pdf>
- [2] PSpice – includes PSpice A/D, PSpice A/D Basics, and PSpice, User's Guide, dostupné na internete, [http://www.seas.upenn.edu/~jan/spice/PSpice\\_UserguideOrCAD.pdf](http://www.seas.upenn.edu/~jan/spice/PSpice_UserguideOrCAD.pdf)
- [3] Katalógový list IRF3205, dostupné na internete, <http://www.irf.com/product-info/datasheets/data/irf3205spbf.pdf>
- [4] Configuration of PSpice Model Libraries, Application note, Author/Date: Wei Ling / 24.08.2010, dostupné na internete, [http://www.flowcad.de/AN/FlowCAD\\_AN\\_PSpice\\_Model\\_Config.pdf](http://www.flowcad.de/AN/FlowCAD_AN_PSpice_Model_Config.pdf)
- [5] The Model Editor – Training video – PSpice A/D, <http://www.youtube.com/watch?v=PPAKEe0XHqo>

# Comparison of Feature Extraction Algorithms for the Acoustic Event Detection System

*Eva KIKTOVÁ, Jozef JUHÁR*

Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic  
eva.kiktova@tuke.sk, jozef.juhar@tuke.sk

**Abstract** – This paper brings the comparison of different feature extraction algorithms for the acoustic event detection system. Many attentions were devoted to the front-end processing where acoustic signals are analyzed in a temporal, spectral and cepstral domain. This work provides also an investigation of appropriate settings of input parameters for considered extraction algorithms. The Hidden Markov Model classification technique was successfully applied and the best recognition performance of each extraction algorithm is presented at the end of this experimental work.

**Keywords** – Acoustic event, feature extraction, detection system

## I. INTRODUCTION

An acoustic signal analysis and classification represents a way how to describe human activity or environmental conditions. Each sound belongs to some sound source, thus the detection of the particular sound event determines the type of the sound source. For example footsteps, sounds of human falls or cry appointed to the presence of human and also it is possible to estimate a particular situation or an activity (normal or abnormal). An examples of such solutions can be found in works [1],[2].

In this work we analyze two types of events, i.e. gunshots and breaking glass. Mentioned sounds are called like acoustic events or foreground sounds. They can be characterized such as unpredictable, rare and relative short sounds, which have a high relevance. These two sounds and many other (e.g. explosions, sounds of road accidents, vulgarisms, etc.) can point to the dangerous situations and their detection can be helpful, e.g. for the police. The system that is able to detect abnormal sounds and generates alarm messages uses abbreviation AED (Acoustic Event Detection) system.

The realization of such system includes a phase of design, a definition of suitable model prototypes and a realization phase that includes also end-users specifications. The AED system may follow two basic operation strategies i.e. "detection and classification strategy" and "detection by classification strategy". Our system solution EAR-TUKE [3] follows the detection by classification strategy, which is characterized by a more complex background modeling, where background model covers all background sounds with varying SNR (Signal Noise Ratio). The proposed system doesn't use VAD (Voice Activity Detection) algorithm, which is applied in the first AED operation strategy. EAR-TUKE continuously monitors and classifies all sound samples.

Each AED system consists of two main functional blocks. First block is responsible for the acoustic signal processing and the feature processing (extraction, transformation, selection, etc.). This block is also called like a front-end processing block. The core of the second block is a classifier, in which input samples are divided to predefined categories. In this work, Hidden Markov Model (HMM) approach was successfully employed. This paper is devoted mainly to the front-end processing, especially to the feature extraction problem. Many feature extraction approaches were investigated and some of them are also employed in other applications, e.g. Mel-Frequency Cepstral Coefficients (MFCC) that have very good performance in tasks related to the speech such as speech or speaker recognition and they are also employed in non-speech tasks such as music retrieval, acoustic event recognition task, etc. [4], [5]. Other group of features was proposed in the MPEG-7 standard [6], where 17 low level descriptors were defined. Some of them were precisely investigated in this paper (mainly basic and spectral descriptors). A special attention was devoted to the appropriate input variables such as frequency range, time step and



spectral resolution. Each type of feature was used to train HMM classifier, which was employed in the testing process using Viterbi based decoding algorithm.

The remainder of this paper is organized as follows. Section II describes the theory of extracted features. Section III includes the practical information about the feature computation. Section IV is devoted to system description and obtained experimental results. The concluding remarks of this work are presented in the last Section V.

## II. FEATURE EXTRACTION - THEORETICAL BACKGROUND

Some features employed in this work were inspired by the speech recognition domain but they are also able to recognize the acoustic events (shot and breaking glass). In the next lines the theoretical background of used extraction approaches are briefly described.

### A. Mel-filter bank coefficients (MELSPEC), Log-frequency filter bank coefficients (FBANK), Mel-frequency cepstral coefficients (MFCC) and Power Normalized Cepstral Coefficients (PNCC)

The computation of MELSPEC, FBANK, MFCC features [6], [7], is following: a pre-emphasis filtration, then an audio signal is divided into the short frames, where the parameters of the signal are constant, then a Hamming window is applied. Signal frames are transformed to the frequency domain via the discrete Fast Fourier Transform (FFT), and then the magnitude spectrum is passed through a bank of triangular shaped filters. MELSPEC features are the output of mentioned operations. The energy output from each filter is then log - compressed, so FBANK features represent the log-frequency filter bank coefficients. The application of the discrete cosine transform (DCT) represents the last step in the MFCC computation. FBANK and MELSPEC features are strongly inspired to the classical MFC (Mel-Frequency Cepstral) analysis.

Some works refer to Power Normalized Cepstral Coefficients (PNCC) [8] like a more robust in the presence of various types of additional noises and in reverberant environments. The PNCC processing is similar to conventional MFCC processing. PNCC computation includes the Gammatone filtration, the use of a power-law nonlinearity that replaces the traditional log operation used in MFCC, a noise-suppression algorithm based on the asymmetric filtering that suppress the background excitation, a module of temporal masking and normalizations blocks. The final PNCC features are decorrelated via DCT transform.

### B. Linear prediction coefficients (LPC), Linear prediction cepstral coefficients (LPCC) and Perceptual linear prediction (PLP) coefficients

Linear prediction analysis of speech is historically one of the most important speech analysis techniques [7], [9]. LPC parameters are widely used in various applications. They represent the short-time spectral envelope information of the signal. LPC is based on a predictive analysis assuming that a sample of signal at the current time is a linear combination of past n-signal samples. The basis is the source-filter model where the filter is constrained to be an all-pole linear filter. The LPC coefficients of each frame are found by applying Levinson-Durbin algorithm.

LPCC coefficients [10] have their background in the LPC coefficients. These are the coefficients of an auto-regressive model. Cepstrum is defined as coefficients of the Inverse Fourier Transform (IFT) representation of the log magnitude of the spectrum. Cepstral analysis is used to separate the source from the coefficients of the vocal tract filter. In the speech recognition, LPCC coefficients prove to capture most of the relevant information contained in the original speech signal.

PLP features are identical to LPC features except that their spectral characteristics have been transformed to match characteristics of human auditory system. PLP approximates three main perceptual aspects namely: the critical-band resolution curves, the equal-loudness curve, and the intensity-loudness power-law relation. PLP are often used as an alternative to the MFCC coefficients.

### C. Spectral features: Audio Spectrum Envelope (ASE), Audio Spectrum Centroid (ASC), Audio Spectrum Spread (ASS), Audio Spectrum Flatness (ASF), Spectral Flux and Spectral Roll-Off

The Audio Spectrum Envelope (ASE) [6] is obtained by summing the energy of the power spectrum within frequency bands which are logarithmically distributed between lower and higher frequencies. The sum of power coefficients in defined band gives the ASE coefficient for this frequency range. The ASE provides a compact representation of the spectrogram of the input acoustic signal.

The Audio Spectrum Centroid (ASC) gives information about the centre of gravity of a log-frequency power spectrum [6]. The ASC measure describes the shape of the power spectrum



and indicates whether in a power spectrum low or high frequencies are dominated. It is an approximation of the perceptual sharpness of the signal.

The Audio Spectrum Spread (ASS) is a measure of the spectral shape and it is defined as the second central moment of the log-frequency spectrum [6]. ASS is extracted by taking the root-mean-square (RMS) deviation of the spectrum from its centroid ASC. The ASS gives indications about how the spectrum is distributed around its centroid. A low ASS values indicated that the spectrum is concentrated around the centroid, where as a high values reflect a distribution of power across a wider range of frequencies [6].

The Audio Spectrum Flatness (ASF) reflects the flatness properties of the power spectrum [6]. For a given signal frame, it consists of a series of values, each one expressing the deviation of the signals power spectrum from a flat shape inside a predefined frequency band. For each band, a spectral flatness is estimated as the ratio between the geometric mean and the arithmetic mean of the spectral power coefficients within this band. High ASF coefficients reflect noisiness whereas a low values may indicate a harmonic structure of the spectrum.

Spectral Flux [6] is a measure of how quickly the power spectrum of a signal is changing. It is a measure of the variation of the power spectrum between adjacent frames. Spectral Flux can be used to determine the timbre of an audio signal.

Spectral Roll-Off frequency [6] is a measure defining the frequency below which 85% of the accumulated magnitude of the spectrum is resides. The threshold takes values between 0.85 and 0.99. The Roll-Off reflects the "skewness" of the spectral shape.

#### D. Temporal features: Audio Waveform (AW), Skewness, Kurtosis and Zero Crossing Rate (ZCR)

Audio Waveform (AW) [6] gives information about minimum and maximum values from the acoustic signal. It is the straightforward description of the shape of any audio signal.

Skewness [11] is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point. The Skewness for a normal distribution is zero and any symmetric data should have a Skewness near zero.

Kurtosis [11] is a measure of the "peakedness" of the probability distribution. The Kurtosis for a standard normal distribution is equal to three.

Zero-Crossing Rate (ZCR) [6] is the rate of how often the signal changes its sign within a signal frame, i.e., the rate at which the signal changes from the negative to the positive or vice versa.

### III. FEATURE COMPUTATION AND EXPERIMENTS

All speech based features except PNCC were extracted in the HTK environment [7]. PNCC, descriptors from MPEG-7 standard and the rest of features were extracted in Matlab environment. The MPEG-7 descriptors (ASE, ASC, ASS, ASF, AW) were extracted with different input parameters e.g. *hopSize*, *loEdge*, *hiEdge*, *octaveResolution*. The rest of spectral features were extracted with the *frame length* and the *overlap* parameters. Time expressing features required only the *frame length* input parameter [5]. The spectrograms of analyzed signals are depicted in the Fig. 1 and the pictures for temporal domain are depicted in the Fig. 2.

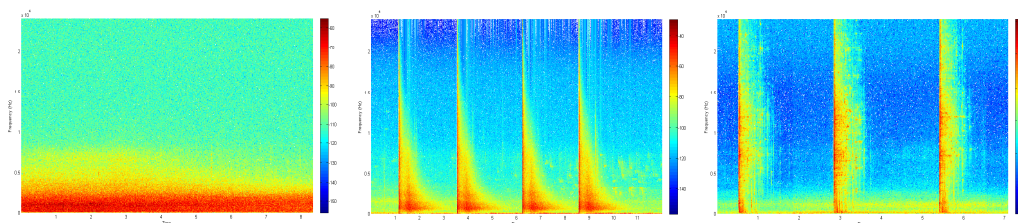


Fig. 1 Spectrograms of background sound, shot and broken glass.

Experiments were done by sound data (48 kHz, 16 bits per sample) from our extended JDAE TUKE database [12]. It contains gun shot recordings with 463 realizations of shots from commonly used weapons, breaking glass recordings with 150 realizations of broken glass and approximately 53 minutes of background sounds (traffic). Recordings were cut and manually labeled. In the testing process 46 gunshots and 13 breaking glass were used. Shots and breaking glass sounds formed one class i.e. the acoustic events class (or foreground sounds).

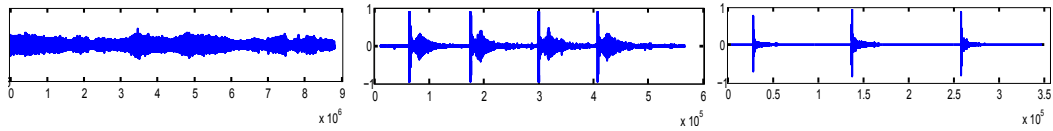


Fig. 2 Temporal characteristics of background sound, shot and broken glass.

#### A. MELSPEC, FBANK, MFCC and PNCC experiments

MFCC coefficients were extracted without pre-emphasis filtration (it is suitable for speech related tasks), with 25 ms Hamming window, 10 ms of frame rate and with 22 mel-filters (12 in the case of MELSPEC). The final number of coefficients has been set to 12. Many other configurations were evaluated, but the optimal results were achieved with presented settings. Their extractions were realized in HTK environment.

According to the mentioned settings, MELSPEC, FBANK and MFCC features were extracted with energy coefficient `_E`, delta and acceleration coefficients `_D_A` [7]. The positive impact of these parameters was confirmed in our internal experiments too. This way, 39 dimension feature vectors were extracted from signal frame.

PNCC were extracted in Matlab environment with 25 ms frames and 10 ms of frame rate with default settings. Feature vectors were extended by the computation of `_D_A` coefficients to the final dimension 39.

#### B. LPC, LPCC and PLP experiments

LPC and LPCC features were extracted with 25 ms Hamming window with 10 ms of frame rate and 12 or 14 order of LPC filter for LPC or LPCC. The final number of coefficients was set to 12. The log-energy (`_E`) feature and `_D_A` coefficients were computed too and finally, the feature vectors were 39 coefficients per frame, i.e. dimension 39.

#### C. Spectral features

1) *ASE experiments* : We are focused on the MPEG-7 low level audio descriptors too. Our attention was concentrated especially on the basic spectral descriptors ASE, ASC, ASS and ASF. Each of them was extracted with different settings.

Many experiments were done with ASE descriptor, which describes the log-frequency power spectrum of the acoustic signal. We investigated the performance of ASE for the detection of acoustic events.

The influence of the frame length was investigated by the `hopSize` parameter, which was set to 10, 20, 30, 50, 100 ms. The best experimental value for `hopSize` parameter was 10 ms. Then several tests for finding the suitable frequency limits for lower edge (`loEdge`) and higher frequency edge (`hiEdge`) were done. According to MPEG-7 recommendation `loEdge` and `hiEdge` should be set to 62.5 Hz and 16000 Hz. The extraction of ASE was also done by the 62.5, 1000, 2000 for `loEdge` and 16000, 22000, 23000 Hz for `hiEdge` frequency limit.

The most suitable `loEdge` was 1000 Hz and `hiEdge` was 23000 Hz. ASE with 1/8 frequency resolution was extracted. Originally, ASE had 38 coefficients per frame (dim 38).

2) *ASF experiments* : Many experiments with ASF descriptor for the purpose of appropriate spectral flatness description were performed. Each change of the input acoustic signal causes the change in ASF coefficients. Generally, ASF descriptor generates  $m$ -dimensional feature vector per signal frame depending on the frequency band, which width is appointed according to the `loEdge` and `hiEdge` parameters. The `loEdge` parameter was set to 250, 1000, 2000, 2500, 4000, 10000 and `hiEdge` was set to 22000. We also investigated the appropriate size of the `hopSize` parameter. ASF with 3, 4, 5, 10, 20, 30 ms of `hopSize` was extracted.

The best results were achieved by ASF features from 2000 Hz to 22000 Hz frequency range and with 4 ms of `hopSize`. The ASF feature vectors had 17 coefficients (dim 17).

3) *ASC and ASS experiments*: ASC describes the centre of gravity of a log-frequency power spectrum and ASS gives information about how the spectrum is distributed around its centroid. ASC and ASS were computed separately from input signal with the same `hopSize`. These descriptors generated one feature per frame. `HopSize` was set to 3, 4, 5, 10, 20 ms. Similar results were reached by the ASC and ASS features extracted with 4 a 5 ms of `hopSize` but the best value was 4 ms.

4) *Flux and Roll-off experiments* : Similarly, spectral Flux and Roll-Off were extracted with the frame length and overlap. They generated one coefficient per frame (dim 1). Spectral Flux is a measure how quickly the spectrum of the signal is changing. Appropriate length and overlap for Flux were 25/10 ms and 10/5 ms and the best results were reached for 10 ms frames and 5 ms overlap.

Spectral Roll-Off coefficients express the skewedness of the spectral distribution. Roll-off coefficients extracted with 10 ms length frame and 5 ms overlap achieved the best results.

#### D. Temporal features

1) *AW experiments*: AW describes signal by two parameters (dim 2), i.e. the minimum and maximum value of the signal frame. Features were extracted with *hopSize* = 3, 4, 5, 10, 20 ms. The best results were achieved with *hopSize* = 3.

2) *ZCR experiments*: Parameter ZCR describes the signal changes from positive to negative or vice-versa. Different settings of frame length were evaluated (4, 5, 10, 20 ms) and the best results were reached with 5 ms of frame length.

3) *Skewness and Kurtosis experiments*: Statistical moments such as Skewness and Kurtosis were extracted with frame length equal to 5, 10, 20, 25, 30 ms. The suitable length was 10 ms for Skewness and 25 ms for Kurtosis parameter.

### IV. RESULTS OF EXPERIMENTS FOR AED SYSTEM

Performances of individual features are presented here. HMM models from one to four states and from 1 to 1024 Probability Density Functions (PDFs) were employed like the classifier. In the evaluation phase, the word insertion penalty -100 [7] were applied. Obtained results were expressed by two measures i.e. ACC[%] and F-measure[%].

*Accuracy* measure is defined as follows:

$$ACC [\%] = \frac{N - D - S - I}{N} \times 100, \quad (1)$$

where  $D$  is the number of deletion errors,  $S$  the number of substitution errors,  $I$  the number of insertion errors, and  $N$  is the total number of labels in the reference transcription files [7]. It informs about the overall system performance.

Another evaluation approach is oriented on the particular class (acoustic events) and several measures can be computed e.g. *Precision*, *Recall*, *F - measure*, etc.

*F-measure* is defined as follows:

$$F - measure = \frac{(\beta_2 + 1) \times Precision \times Recall}{\beta_2 \times Precision + Recall}, \quad (2)$$

where *Precision* and *Recall* are computed by the following formulas:  $Precision[\%] = \frac{TP}{TP+FP} \times 100$  and  $Recall [\%] = \frac{TP}{TP+FN} \times 100$ .

$TP$  - true positive is the number of items correctly labeled as belonging to the positive class,  $FP$  - false positive is the number of items which are incorrectly labeled as belonging to the positive class,  $FN$  - false negative corresponds to the items which were not labeled as belonging to the positive class but should have been. *F-measure* combines precision and recall measures and it is balanced when  $\beta_2 = 1$ . For each type of features, the best score was depicted in the Fig. 3.

The best results were achieved for speech inspired features, especially in the case of MELSPEC features, where some models achieved perfect recognition results 100% of ACC and F-measure. Then follow FBANK and MFCC. The visualization of mentioned features is depicted in the Fig.4.

The best one from MPEG-7 spectral descriptors was ASE (dim 38) that achieved score up to 80% of ACC and F-measure. From the last group, i.e. temporal based features, the best score was achieved by AW descriptor (dim 2) that achieved score up to 70% of ACC and F-measure. On the other hand the individual performance of the Skewness, Spectral flux and PNCC features seems to be less effective for the acoustic event detection task.

### V. CONCLUSION

In this experimental work, the comparative study of individual features was presented. Applied feature extraction algorithms described acoustic signals by different ways, where some obtained features demonstrated a very powerful discriminative ability and highlighted desired characteristics of acoustic events. Some approaches were not so effective, especially in the PNCC case, where some operations that ensure the robustness of this approach in speech related tasks caused the degradation of recognition results in the acoustic events detection task. Low recognition performances of other scalar features could be expected, because only one parameter expressed the signal characteristic. It could affect also the process of HMM training.

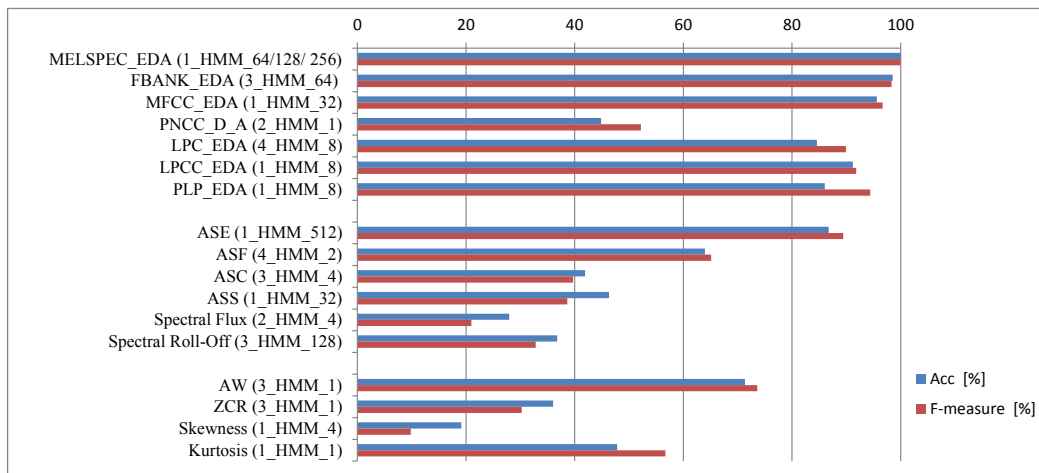


Fig. 3 Results of experiments. (The name composed of the feature's kind and type of HMM (state\_HMM\_PDFs)).

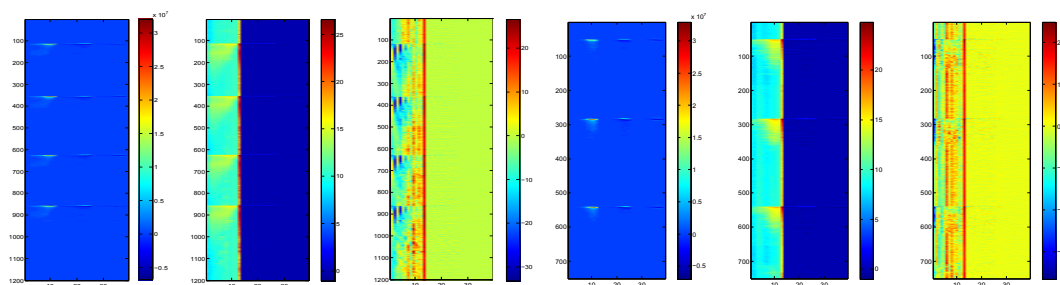


Fig. 4 Visualization of MELSPEC, FBANK and MFCC coefficients (first 3 ones for shot, last 3 for glass events).

#### ACKNOWLEDGMENT

This research work was supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project VEGA 1/0386/12 (30%) and the Research & Development Operational Programme funded by the ERDF under the project UVP Technicom ITMS-26220220182 (70%).

#### REFERENCES

- [1] Y. Zhan, S. Miura, J. Nishimura, and T. Kuroda, "Human activity recognition from environmental background sounds for wireless sensor networks," in *Networking, Sensing and Control, 2007 IEEE International Conference on*, April 2007, pp. 307–312.
- [2] Y. Li, Z. Zeng, M. Popescu, and K. C. Ho, "Acoustic fall detection using a circular microphone array," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, Aug 2010, pp. 2242–2245.
- [3] M. Lojka, M. Pleva, E. Kiktova, J. Juhar, and A. Cizmar, "Ear-tuke: The acoustic event detection system," in *Multimedia Communications, Services and Security*, ser. Communications in Computer and Information Science, 2014, vol. 429, pp. 137–148.
- [4] C. Clavel, T. Ehrette, and G. Richard, "Events detection for an audio-based surveillance system," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, July 2005, pp. 1306–1309.
- [5] E. Kiktova-Vozarikova, J. Juhar, and A. Cizmar, "Feature selection for acoustic events detection," *Multimedia Tools and Applications*, pp. 1–21, 2013.
- [6] H.-G. Kim, N. Moreau, and T. Sikora, *MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval*. Wiley, Oct. 2005.
- [7] S. J. Young, G. Evermann, M. J. F. Gales, T. Hain, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. C. Woodland, *The HTK Book, version 3.4*. Cambridge, UK: Cambridge University Engineering Department, 2006.
- [8] C. Kim and R. Stern, "Power-normalized cepstral coefficients (pncc) for robust speech recognition," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, March 2012, pp. 4101–4104.
- [9] N. Morgan, H. Hermansky, H. Bourlard, P. Kohn, and C. Wooters, "Continuous speech recognition using plp analysis with multilayer perceptrons," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, Apr 1991, pp. 49–52 vol.1.
- [10] G. Antoniol, V. F. Rollo, and G. Venturi, "Linear predictive coding and cepstrum coefficients for mining time variant information from software repositories," *SIGSOFT Softw. Eng. Notes*, vol. 30, no. 4, pp. 1–5, May 2005.
- [11] "Engineering statistics handbook." [Online]. Available: <http://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm>
- [12] M. Pleva, E. Vozarikova, L. Dobos, and A. Cizmar, "The joint database of audio events and backgrounds for monitoring of urban areas," *Journal of Electrical and Electronics Engineering*, vol. 4, no. 1, pp. 185–188, 2011.

# Inteligentná elektrická inštalácia CyBrotech

<sup>1</sup>Štefan ŠUTOR, <sup>2</sup>Stanislav ILENIN

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
univerzita v Košiciach, Slovenská republika

<sup>1</sup>stefan.sutor@student.tuke.sk, <sup>2</sup>stanislav.ilenin@tuke.sk

**Abstrakt** — Tento článok sa zaoberá problémom automatizácie budov systémom CyBrotech. Rieši možnosti návrhu elektrickej inštalácie rodinného domu. Inteligentná inštalácia využíva všetky prvky moderných technológií a kombinácie IT prvkov. Vďaka tomu nám CyBrotech poskytuje možnosť jednoduchej a cenovo dostupnej inteligentnej elektroinštalácie.

**Kľúčové slová** — inteligentná elektroinštalácia, CyBrotech, zbernicový systém

## I. ÚVOD

V dnešnej dobe sa čoraz viac začínajú používať pojmy nízkoenergetický dom a inteligentná budova. Pod pojmom inteligentná budova môžeme rozumieť objekt, ktorý je plne automatizovaný a dokáže sa sám riadiť na základe vopred nastavených podnetov a informácií zo senzorov. Nízkoenergetický dom je definovaný ako budova, ktorej spotreba tepla na vykurovanie nepresiahne hodnotu 50 kWh/m<sup>2</sup>. Systémy inteligentných elektroinštalácií prestávajú byť prednosťou veľkých stavieb a začínajú sa pomaly používať aj v malých rodinných domoch, kde ponúkajú svojim užívateľom nie len funkčnosť, ale aj komfort a bezpečnosť. Firma CybroTech v roku 2005 prišla na trh so systémom Integra IQ @ HOME, Integra IQ @ HOTEL a Integra IQ @ OFFICE. Integra IQ @ HOME je určený pre novostavby, rekonštruované byty a rodinné domy [1]. Je to systém automatizácie domácnosti (často nazývaný aj inteligentná domácnosť). Systém Integra IQ @ Home zahŕňa kontrolu všetkých štandardných systémov zahrnutých v každej modernej domácnosti, ako sú: riadenie osvetlenia, regulácia teploty, ovládanie žalúzií, alarm, a detektory úniku plynu alebo úniku vody. Systém zahŕňa aj IP videotelefony, pomocou ktorých môžeme prakticky zo všetkých inteligentných zariadení v dome (iPhone, iPad, Android, atď) riadiť všetky ovládané zariadenia. Inštalácia a konfigurácia systému je jednoduchá a veľmi rýchla bez programovania. Integra IQ @ HOME využíva výhody moderných terminálov (dotykový panel, smartfóny, tablety...), ktoré umožňujú zjednodušenie každodenného života [2].

## II. PRVKY INTELIGENTNEJ ELEKTROINŠTALÁCIE CYBRO

### A. Riadiace moduly

*Riadiaci modul CyBro-2-230.* Riadiaci modul je základný vysokovýkonný, programovateľný prvok systémovej elektroinštalácie, ktorý riadi komunikáciu prvkov primárnej zbernice IEX BUS. Komunikácia s PC je realizovaná cez jednoduché pripojenie sériovej (RS232) linky cez RJ9(sériový) konektor alebo Ethernet linky cez RJ45 (Ethernet) konektor. Taktiež môže byť prepojený pomocou Ethernet portu s ďalšími systémami automatizácie. Riadiaci modul je vybavený LED ukazovateľmi pre riadiace a komunikačné stavy 10 digitálnymi vstupmi a 8 digitálnymi výstupmi. Napájanie je realizované zo siete 230 V AC čo je jeho veľkou výhodou, pretože nepotrebuje napájací modul, ktorý znamená ďalšiu investíciu a taktiež zabratie miesta v rozvážači. Obsahuje napäťový menič, ktorého vstupné napätie je 230V AC a výstupné napätie je 24V DC [3].

*Riadiaci modul RC (room controller).* Riadiaci modul RC je základný vysokovýkonný a programovateľný prvok, ale neobsahuje žiadne digitálne ani analógové vstupy a výstupy.

Nastavenia sa realizujú programom CyBro integrovaného vývojového nástroja cez osobný počítač alebo operačným panelom s LCD dotykovým displejom. Komunikácia medzi rozširujúcimi modulmi a ostatnými prvkami prebieha rovnako ako pri CyBro-2-230 pomocou IEX-2 zbernice a cez sériové linky. Na komunikáciu s ostatnými riadiacimi jednotkami sa využíva sieť A-bus alebo ethernetová sieť. Už z jeho názvu „izbový regulátor“ je možné povedať, že sa jedná o riadiaci prvok, ktorý sa používa hlavne v obývacích izbách, spálňach alebo detských izbách [4].

#### B. Moduly umiestnené v rozvádzači

*Bio 24R.* Bio-24R je IEX rozširujúci modul s dvanástimi dvojsmernými vstupmi a dvanástimi reléovými výstupmi. Reléové výstupy môžu byť použité ku kontrole hlavných energetických zariadení ako sú napríklad svetlá, žalúzie, rolety a mnoho iných. Stav individuálnych vstupov, výstupov a komunikácie je indikovaný LED diódami. IEX porty umiestnené bočne na puzdre umožňujú ľahkú rozšíriteľnosť o ďalšie moduly.

*HR (hotel room controller).* HR je jednoduchý a rentabilný riadiaci modul izieb. Kontroluje vykurovanie, chladenie, ventiláciu, osvetlenie a požiarne alarm. 12 digitálnych vstupov sa môže použiť pre vypínače alebo termostaty. 10 reléových výstupov môže kontrolovať jednotku fan-coil, osvetlenie a pod. Dva externé senzory môžu byť pripojené k meraniu izbovej teploty. IEX porty ležia na stranách modulu, ktoré slúžia aj na napájanie modulu.

*GSM-1.* GSM/GPRS modul je použitý na kontrolu systému automatizácie. Užívateľ môže prijímať SMS správy so systémovými informáciami alebo posielat' SMS správy s príkazmi. Obsahuje 4 analógové/digitálne vstupy a dva reléové výstupy. Z jeho zadnej strany sa nasúva SIM karta. Ak jednotka je umiestnená mimo rozvádzač, potrebuje podporu napájacieho napätia 24V DC.

#### C. Moduly umiestnené v miestnosti

*FC.* FC môže byť použitý ako riadiaci modul pre fan-coil jednotku s tromi relé, z ktorých jedno kontroluje rýchlosť ventilátora a ostatné dve relé sú pre ventily teplej a studenej vody. Jednotka obsahuje jeden digitálny vstup, ktorý môže byť použitý ako vstup signálu z okenného kontaktu (otvorené alebo zatvorené), a tak dá signál automatizačnému systému k zníženiu alebo úplnému vypnutiu klimatizácie. Obsahuje dva porty pre pripojenie externých tepelných snímačov na zistenie teploty miestnosti. Jednotka FC je chránená plastovou montážnou škatuľou s triedou ochrany IP 42, ktorá sa montuje buď priamo v zariadení fan-coil, na stene alebo na DIN lište. Napájaná je napätím 24V DC z IEX zbernice [5].

*LC-S.* LC-S je modul na kontrolu osvetlenia so vstupmi využívanými v svetelnej automatizácii. Výstupy 0...10 V umožňujú pripojiť analógové predradníky a ovládať stmievanie osvetlenia. Pohybové senzory aj multisenzory môžu byť pripojené k modulu FC k poskytovaniu informácií ohľadom obsadenia a osvetlenia miestnosti. Tri analógové vstupy sú priradené k stmievacím potenciometrom. Modul je chránený plastovou škatuľou so stupňom ochrany IP42. Môže byť inštalovaný v podlahe, na strope alebo na DIN lište. Napájanie je zabezpečené zbernicou IEX-2.

#### D. Senzory

*ES-W.* ES-W je externý teplotný senzor, ktorý môže byť pripojený so systémom automatizácie cez TS-H (senzor vlhkosti a teploty), FC alebo HR (hotel room controller) IEX moduly. Senzor ES-W umožňuje presné meranie teploty pomocou káblového senzora s veľmi malými odchýlkami. Maximálna vzdialenosť od TS po ES musí byť maximálne 3 metre. Vzdialenosť od iných zariadení podporujúcich tento senzor je maximálne 20 metrov, maximálna pracovná odchýlka je  $\pm 0,5^{\circ}\text{C}$  ( $-10^{\circ}\text{C} \div +85^{\circ}\text{C}$ ) [6].

*MS.* MS (multisenzor) je snímač osvetlenia pre aktuálne osvetlenie v miestnosti, automatizačný stupeň osvetlenia možno nastaviť na požadovanú úroveň. MS možno pripojiť k systému automatizácie cez LC-S modul (regulátor osvetlenia). MS obsahuje taktiež IR prijímač, ktorý zisťuje príkazy z diaľkového ovládania a umožňuje tak užívateľovi ovládať systémové parametre. Pri inštalovaní si musíme uvedomiť, ktorú časť bude MS sledovať pretože poskytuje možnosť rôznych snímacích uhlov. Napríklad pri sledovaní vonkajšieho svetla snímačom umiestneným v objekte sa z meraní zistilo, že snímač má najlepšiu citlivosť do  $45^{\circ}$ . Pri sledovaní vnútorného osvetlenia MS ponúka tri varianty snímania. Pracovný uhol v horizontálnej rovine môže byť  $60^{\circ}$  alebo aj  $360^{\circ}$ , to závisí na umiestnení interného snímača. Vo vertikálnej polohe nám snímač ponúka operačný uhol  $150^{\circ}$ . Diaľkový IR ovládač je možné použiť do 5 m.



*LRI8134.* Multisenzor LRI8134 je kompaktná jednotka obsahujúca tri odlišné senzory požadované k poskytovaniu komfortu a výhod úspory elektrickej energie kontrolou svetelného systému. Obsahuje svetelný senzor, PIR pohybový senzor a IR prijímací modul pre manuálne ovládanie pomocou IR diaľkového ovládania. Táto jednotka je navrhnutá pre vnútorné použitie k poskytovaniu informácií obsadenosti miestnosti a osvetlenia. Vyhodenie je optimalizované pre umiestnenie na strop vo výške medzi 2,4 a 3,5 m nad podlahou. Multisenzor je napájaný z LC modulu 24 V DC cez zbernicu RJ12 konektora.

*LRM8117/00.* Pohybový PIR mikro detektor navrhnutý pre vnútorné prostredie. Integrovaný otočný držiak umožňuje aby hlava senzora sledovala daný priestor horizontálne aj vertikálne. Senzor je vybavený možnosťou potlačenia denného svetla. V prípade dostatočného denného svetla bude senzor držať výstupný signál (svetlá ostanú vypnuté), kým prirodzené svetlo neklesne pod referenčnú hodnotu osvetlenia danú užívateľom. Táto funkcia je továrensky nastavená. Najvýhodnejšia inštalácia je v priestoroch, kde obsadenosť miestnosti je riedka alebo nepredvídateľná. Jeho podpora dlhého dosahu snímania pomocou freaselovej šošovky, je ideálna pre použitie v úzkych priestoroch.

*THS02.* THS02 je mechanický izbový termostat kontrolujúci vykurovanie, chladenie a taktiež má možnosť kontrolovať otáčky ventilátora. Obsahuje integrovaný digitálny senzor, ktorý umožňuje presné meranie teploty v miestnosti s malými odchýlkami. THS môže byť ľahko pripojený k digitálnym a teplotným vstupom HR modulu.

*OP-MT8070i.* MT8000 séria je nová HMI generácia Wientek-u. Panelové pripojenie k riadiacej jednotke CyBro-2 je umožnené cez sériovú a Ethernetovú komunikáciu. Produktový rad ponúka jednotky s rôznymi rozmermi, farbami displeja a dostupnej pamäte. Každý panel používa odporovú technológiu detekcie dotyku, čo umožňuje vytváranie výhod moderných rozhraní človek – stroj. Operačný panel je napájaný s externého zdroja pomocného napájania 24 V DC. OP-MT8070i má 7" displej s rozlíšením 800x480, čo nám ponúka pekné farebné prevedenie pracovného prostredia OP. Prostredie nám ponúka prehľad všetkých inštalovaných prvkov použitých v dome a taktiež je možnosť si nastaviť pohľad na konkrétnu izbu alebo celý dom (byt). OP obsahuje 128 MB používateľskej pamäte, 1x čítačku kariet a USB port. USB port a čítačka kariet umožňujú užívateľovi obnovenie dát pri ich strate, preprogramovanie alebo zadať nové nastavenia funkcií bez nutnosti použiť osobný počítač. Predná časť operačného panela je veľmi dobre chránená, pre ktorú výrobca udáva stupeň ochrany IP65.

### III. PROJEKT INTELIGENTNEJ ELEKTROINŠTALÁCIE CYBRO

#### A. Základné obvody a zariadenia

Projekt rieši návrh inteligentnej elektrickej inštalácie rodinného domu. Prípojka je navrhnutá káblom AYKY-J 4x16mm<sup>2</sup> z káblového podzemného vedenia napojená zaslučkovaním. Ukončenie prípojky je v prípojkevej skrini, ktorá spolu s elektromerovým rozvádzačom je umiestnená v oplotení. Z elektromerového rozvádzača je napájaný hlavný rozvádzač, umiestnený v garáži. Z hlavného rozvádzača sú napájané niektoré silové obvody a dva podružné rozvádzače osadené prvkami CyBro. Jeden podružný rozvádzač je umiestnený v garáži a druhý v obývacej izbe.

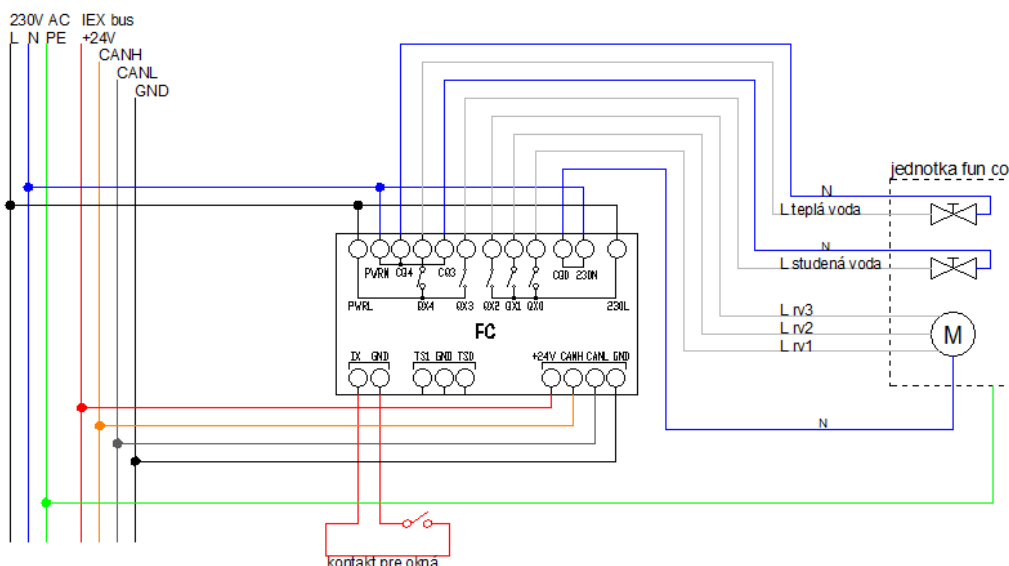
Napájacie napätie musí byť privedené do prvého zariadenia v rozvádzači. Ak sú zariadenia po prvý krát napojené na zdroj elektrického napätia, automatické adresovanie zariadení začne automaticky. Ak sú zariadenia premiestnené alebo zmenili poradie príkaz na auto adresovanie musí byť vyslané zo zariadenia (60s stlačenie) alebo zo serveru mini scada programovým príkazom.

Zariadenia v podružnom rozvádzači sú usporiadané postupne, zľava doprava. Zariadenia mimo rozvádzača sú usporiadané sériovo vzostupne t.j. najmenšie číslo zariadenia sa priradí prvej adrese, druhé najmenšie číslo druhej adrese atď.

Vo vnútri rozvádzača je prevedenie zbernice 4x flat káblom alebo RJ9 konektorom. Vonkajšie zariadenia sú prepojené zbernicou vyhotovenou netieneným krúteným párovým káblom alebo push-wire terminálom

*Vykurovanie a klimatizácia.* Vo väčšine miestností bude inštalované elektrické podlahové vykurovanie. V týchto miestnostiach budú inštalované otočné termostaty CyBro THS02 alebo operačné panely s možnosťou ovládať aj iné funkcie ako napríklad osvetlenie a pod. K týmto zariadeniam budú pripojené tepelné snímače, ktoré budú zisťovať teploty v jednotlivých izbách, ale aj teplotu vonkajšieho prostredia. Termostaty a tep. snímače budú umiestnené v blízkosti dverí vo výške 1,5 m od podlahy. Dom bude rozdelený na 9 vykurovacích segmentov.

V obývacích miestnostiach je použitá aj jednotka fan-coil, ktorá slúži ako na vykurovanie tak aj na chladenie resp. klimatizáciu.



Obr. 1 Zapojenie riadiacej jednotky FC s riadenou vykurovacou jednotkou fun coil

**Osvetlenie.** Väčšina svietidiel objektu je ovládaná tlačidlovými spínačmi, kde stlačenie horného tlačidla znamená povel na zapnutie a dolného povel na vypnutie. V prípade, že svietidlo bude spínané cez stmievač – stropné svietidla v spálni a detských izbách, krátke stlačenie horného tlačidla znamená rozsvietiť na 100% a krátke stlačenie dolného úplne zhasni. V chodbe, vstupnej hale a v komore je spínanie riešené bezdotykovo pomocou inštalovaných PIR senzorov. V kúpeľni, vo WC a v kotolni je ovládanie osvetlenia riešené klasickými vypínačmi. Pri každom vchode do domu je umiestnené hlavné ovládanie osvetlenia, ktoré pri odchádzaní z domu umožní vypnutie svetla v celom dome. Pri hlavnom vstupe do domu a pri vstupe zo záhrady je umiestnený pohybový senzor do vonkajšieho prostredia, ktorý zapne osvetlenie pri týchto vstupoch.

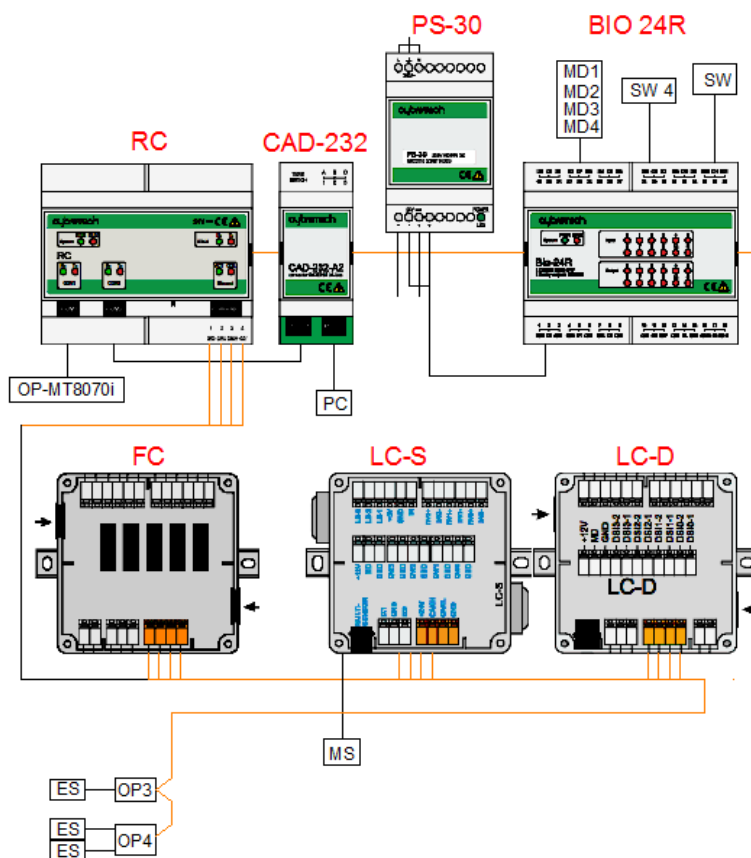
### B. Inštalácia v obývacej izbe

Obývacia izba by mala obsahovať čo najviac funkcií zaručujúcich pohodlie a komfort. Riadiaca jednotka CyBro-2-230 je umiestnená v podružnom rozvážači v garáži. Je výhodnejšie inštalovať druhý podružný rozvážač priamo v obývacej izbe, v ktorom sa nachádza riadiaca jednotka RC len pre túto miestnosť. Komunikácia medzi riadiacimi jednotkami bude zaručená ethernetovou linkou alebo RS232 linkou. Napájanie riadiacich jednotiek, rozširujúcich modulov alebo riadených jednotiek sa nemusí zhodovať so sieťovým napätím 230V AC. Na transformáciu napätia zo sieťového 230 V AC na jednosmerné sú použité PS (power suply) jednotky. Tieto jednotky nenapájajú len riadiace jednotky alebo rozširujúce moduly, ale aj senzory, diódové svietidlá, motorčeky žalúzií a pod.

Hlavným prvkom, ktorý ovláda osvetlenie je rozširujúci modul Bio-24-R. Jeho reléové výstupy riadia činnosť osvetlenia a žalúzií. Bio-24R obsahuje 12 výstupov, z ktorý je polovica výstupov jednosmerného napätia galvanicky oddelená od šiestich výstupov striedavého napätia. Obdobne sú galvanicky rozdelené aj vstupy do rozširujúceho modulu. Bio 24R potrebuje na striedavú časť priviesť 230 V AC a na jednosmernú 24V DC. Jednosmerná časť je napájaná z jednotky PS-30 a riadi motorčeky žalúzií a osvetlenie resp. riadi svetelné efekty a to zmenou farby podsvietenia nábytku. Zmenu farby podsvietenia nám zaručuje jednotka LC-D, z ktorej smerujú 2 výstupy do Bio-24R. Jeden výstup znamená „+“ (otoč doprava/pridaj) a druhý „-“ (otoč doľava/uber). Zmena farby svietidla sa uskutočňuje otáčaním farebnej diskety vo svietidle alebo pridávaním (uberaním) číselných hodnôt (0...255) farieb svetla. Riadenie motorčekov žalúzií je uskutočnené ručne tlačidlom SW-L, na diaľku diaľkovým ovládačom RE alebo automaticky podľa merania osvetlenia multisenzorom MS1 alebo externým svetelným senzorom.

Striedavá časť je pripojená priamo na jednu z fáz elektrickej siete a riadi len osvetlenie. Na ovládanie osvetlenia je postačujúce použiť len tlačidlo SW-L alebo pohybové snímače MD. Ale aby to bolo trochu komfortnejšie, tak na ovládanie osvetlenia je možné taktiež použiť diaľkové ovládanie RE, ktorým bude možné meniť intenzitu osvetlenia, to znamená, že bude poskytnutá možnosť užívateľovi ľubovoľne stlmiť svetlo podľa jeho požiadaviek. Ďalšou z možností je namiesto tlačidla SW-L použiť k zmene intenzity osvetlenia potenciometer. Vtedy sa do jednotky LC-S neprivedie digitálny signál ale analógový.





Obr. 2 Komunikácia medzi prvkami elektroinštalácie v obývacej izbe

Jednotky LC-S a LC-D obsahujú svorky k pripojeniu svetelných multisenzorov prijímačom MS-I/MS-O s IR a jeden káblový konektor pre multisenzor LRI8134. V obývacej izbe sú dva svetelné senzory, ktoré merajú intenzitu osvetlenia. Inštalácia dvoch senzorov je výhodnejšia z toho dôvodu, že miestnosť, resp. okná do miestnosti sú, orientovaná na dve svetové strany. Tieto senzory komunikujú s dvomi senzormi MS-I určenými do vnútorného prostredia a dvomi multisenzormi MS. V prípade dostatku vonkajšieho svetla sa podľa potreby budú natáčať žalúzie a tým prepúšťať svetlo do miestnosti a naopak. V prípade ak je v miestnosti nedostatok svetla a slnečné žiarenie je nedostatočné, tak sa v miestnosti budú rozsvetovať svietidlá podľa vopred nastavenej hodnoty osvetlenia miestnosti.

#### IV. ZÁVER

Na základe poznatkov o prvkoch systému CyBrotech, je možné urobiť návrh inteligentnej elektroinštalácie v ľubovoľnej budove. Je možné povedať, že spoločnosť CyBrotech ponúka pre elektrotechnický trh jednoduchú aplikáciu inteligentnej elektroinštalácie, ktorá je plne pripravená vyriešiť všetky požiadavky užívateľov. Jedným z nedostatkov CyBrotechu, je absencia snímačov zabezpečujúcich protipožiarnu ochranu. Ďalším nedostatkom je to, že ak nastane výpadok elektrickej energie systém sa musí znovu adresovať. Toto nastane len pri výpadku dlhšom ako 10 sekúnd. Riešením je zálohovanie nastavení na pamäťovú kartu, USB alebo počítač, kedy po takomto výpadku sa znovu obnovia predtým nastavené dáta. Ceny prvkov systému CyBro sú pri porovnaní s ostatnými inteligentnými systémami približne rovnaké. Ak porovnáme ceny inteligentnej elektroinštalácie s klasickou silovou inštaláciou, tak tá inteligentná je 2 až 3-krát drahšia, ale za to poskytuje užívateľovi väčší komfort.

#### LITERATÚRA

- [1] Product catalog [online]. Aktualizované september 2011 [cit 2014-1-23]. Dostupné na intrnete: < <http://www.cybrotech.co.uk/partners/Downloads/Catalog/Product%20Catalog%202011-09%20en.pdf>>
- [2] <http://www.robotina.com/products/smart-green/>
- [3] <http://www.cybrotech.co.uk/index.php?lang=en&module=ctlgc&action=product&id=128>
- [4] <http://www.cybrotech.co.uk/index.php?lang=en&module=ctlgc&action=product&id=130>
- [5] <http://www.brookfieldengineering.com/support/faq.asp>
- [6] <http://www.cybrotech.co.uk/index.php?lang=en&module=ctlgc&action=product&id=152>

# Inteligentný elektroinštalačný systém BTicino

<sup>1</sup>Stanislav JENDRICHOVSKÝ, <sup>2</sup>Stanislav ILENIN

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
univerzita v Košiciach, Slovenská republika

<sup>1</sup>stanislav.jendrichovsky@student.tuke.sk, <sup>2</sup>stanislav.ilenin@tuke.sk

**Abstract** — Tento článok sumarizuje základné poznatky o inteligentných elektroinštaláciách. Je uvedená klasifikácia a charakterizovanie rôznych typov riadiacich systémov, ktorých základom je v rôznej miere využívanie funkcií zberníc. Je urobené porovnanie inteligentných riadiacich systémov s klasickou elektroinštaláciou a podrobná charakteristika jedného vybraného riadiaceho systému BTicino, jeho komponentov a využívaných technológií. Zámerom článku je poukázať na možnosti využitia tohto systému a jeho funkcií v rámci riadenia jednotlivých komponentov stavby.

**Keywords** — inteligentná elektroinštalácia, BTicino, BUS, zbernicový systém

## I. ÚVOD

V dnešnej dobe sa pri výstavbe budov kladie čoraz častejšie veľký dôraz na komfort, bezpečnosť a efektívnosť stavaného objektu. V takom prípade je možné tieto požiadavky riešiť využitím moderných elektroinštalačných systémov. Elektroinštaláciu nazývame inteligentnou, ak je tvorená zo sústavy elektrických rozvodov budov vyznačujúcich sa vysokou mierou automatizácie. Je založená na komunikácii medzi jednotlivými prvkami inštalácie s určitou funkciou za účelom docieľiť čo najvyššiu mieru pohodlia užívateľa. Inteligentná inštalácia má oproti klasickej elektroinštalácii veľa pozitív hlavne v prípade monitorovania systému na diaľku, jej ovládania alebo pri šetrení materiálu pri jej výstavbe.

Podľa požiadavky užívateľa je možné vybrať vhodný inteligentný systém pre domácnosť. Pri správnom výbere inteligentnej inštalácie si musí užívateľ určiť, či chce ovládať jej časti ako osvetlenie, vykurovanie a zabezpečovací systém samostatne, alebo chce všetko riadiť centrálné, naraz. Takáto inteligentná inštalácia dokáže riadiť osvetlenie za pomoci vopred navolených režimov, ktoré je možné podľa požiadavky upravovať a meniť.

Tento článok sa zaoberá návrhom inteligentnej inštalácie značky BTicino, keďže tento systém nie je na Slovensku rozšírený a nebola vypracovaná žiadna práca popisujúca túto inštaláciu.

## II. CHARAKTERISTIKA INTELIGENTNÝCH INŠTALÁCIÍ A ICH POROVNANIE S KLASICKOU ELEKTROINŠTALÁCIU

### A. Klasifikácia a charakteristika riadiacich systémov pre budovy

Inteligentnou elektroinštaláciou môžeme nazvať systém elektroinštalácie vo vnútorných priestoroch s vysokou mierou automatizácie. Riadiace systémy budov vyžadujú rozsiahlu komunikáciu. S cieľom klasifikovať inteligentnú elektroinštaláciu budovy, je úroveň riadiaceho systému budovy odstupňovaná podľa miery prístupu a prepojenia jednotlivých častí systémov použitých v budove [1].

*Malé systémy:* budova je minimálne vybavená systémom automatizácie budov - BAS (Building Automation System). BAS je použitý na riadenie jednoduchých úloh, využíva v podstatnej miere prostriedky klasickej elektroinštalácie, systémy sú používané samostatne, nekomunikujú medzi sebou prostredníctvom LAN.

*Stredne veľké zbernicové systémy:* budovy sú systémom automatizácie budov vybavené minimálne. BAS na riadenie využíva zbernicu alebo iný riadiaci vodič, funkčne rieši len jeden subsystém (napríklad len vykurovanie alebo len osvetlenie). Má jednoduché rozhranie, ktoré komunikuje s ostatnými systémami prostredníctvom základnej siete LAN.

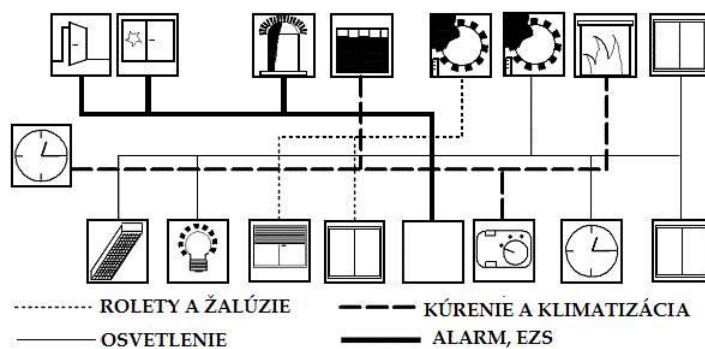
*Komplexná systémová technika:* budovy sú systémom automatizácie budov vybavené vo veľkej miere. BAS používa zbernice na riadenie úloh elektroinštalácie v celej budove, v plnom rozsahu. Spája všetky subsystémy ako osvetlenie, vykurovanie, zabezpečovacie systémy (protipožiarne alebo proti nepovolenému vstupu), spotrebiče na zásuvku a iné, do jedného funkčného celku. Topológia siete LAN je komplexná.

*Distribučná systémová technika:* budova poskytuje najvyššiu úroveň využitia systémov automatizácie budov, jej komunikáciu a riešenia úloh. Aplikuje zložitú sieť LAN. Pre riadenie elektroinštalácie na diaľku a komunikáciu s distribučnými silnoprúdovými sieťami využíva sieť MAN a WAN [2][3].

#### B. Porovnanie zbernicovej a klasickej elektroinštalácie

*Nevýhody klasickej elektroinštalácie:*

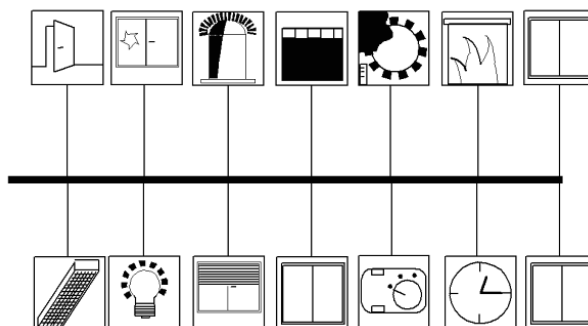
- rozsiahle množstvo vedení, veľkosť systému a dĺžka jednotlivých vedení zvyšuje zaťaženie inštalácie s čím sa spája horšia požiarne bezpečnosť a bezpečnosť osôb,
- veľmi zložitá zapojenie inštalácie,
- malá flexibilita systému (realizácia akejkoľvek zmeny je ťažko realizovateľná),
- vylepšovanie, alebo pridávanie funkcií je veľmi zložitý proces,
- nie je možné realizovať spätné hlásenie, vizualizáciu či centrálné ovládať funkcie inštalácie [4].



Obr. 1 Schéma klasickej elektroinštalácie[5]

*Výhody zbernicovej elektroinštalácie:*

- úspora materiálu, vodičov a energie,
- jednoduchšia manipulácia pri inštalovaní,
- riziko vzájomného nadmerného oteplenia sa zníži v dôsledku zmenšenia počtu vodičov,
- jednoduchá dodatočná rozšíriteľnosť funkčnosti prvkov a možná rekonfigurácia systému,
- jednoduchá diagnostika a servis, vďaka spätnej odozve prvkov, možnosť vizualizácie a počítačového ovládania [4].



Obr. 2 Schéma zbernicovej elektroinštalácie[5]

### III. POPIS INTELIGENTNEJ INŠTALÁCIE BTICINO

Vďaka MY HOME technológii, ktorú firma BTicino používa, je možné navrhovať elektrický systém s vyspelejšími vlastnosťami v nových alebo existujúcich budovách. Systém vytvorený pomocou MY HOME môže byť kedykoľvek rozšírený s minimálnym dopadom na financie a s maximálnou efektívnosťou.

Počas fázy návrhu inštalácie je možné určiť, ktoré funkcie má inštalácia zahŕňať okamžite a ktoré môžu byť doplnené v neskoršom čase. MY HOME je otvorený systém, ktorý dovoľuje integráciu produktov iných značiek a taktiež komunikáciu s vonkajším svetom za pomoci TCP/IP protokolu. Samotná inštalácia a jej návrh je veľmi jednoduchá, vďaka zahrnutiu komunikácie medzi komponentmi digitálnym prenosom pomocou BUS zbernice alebo využitia technológie rádiového prenosu [5].

#### A. Spôsob komunikácie

Technológia SCS BUS je založená na spojení prístrojov pomocou kábla (BUS) s dvomi vodičmi, ktoré slúžia na prenos informácií a napájanie prístrojov nízkym napätím, 27V DC. Použitím vhodných rozhraní, je možné MY HOME systém spojiť s rôznymi komunikačnými protokolmi, ako štandardy Konnex, DALI a pod. a rozšíriť pomocou ZigBee zariadeniami na ovládanie. SCS BUS zapojenie je možné inštalovať pri zapájaní novej inštalácii, alebo pri kompletnej renovácii starej [5].

ZigBee technológia je založená na použití batériou napájaných ovládacích zariadení a akčných členov napájaných z elektrickej siete, ktoré sú pripojené na záťaž, ktorú majú ovládať. Komunikácia prebieha cez rádiový signál na frekvencii 2,4GHz. Vďaka týmto vlastnostiam, je realizácia inteligentného systému veľmi jednoduchá, keďže nie je potrebný zásah do káblovej štruktúry elektrického rozvodu a muriva stien. Použitím správneho rozhrania, je možné ZigBee systém integrovať do existujúceho BUS systému. Systém ZigBee sa inštaluje tam, kde nie je možnosť rozšíriť alebo upravovať BUS kabeláž [5].

#### B. Prvky systému

*Napájací zdroj E46ADCN.* Napájací zdroj je základnou súčasťou MY HOME systému, bez ktorej by inteligentná inštalácia nefungovala. Na výstupe jednotka dodáva nízke striedavé napätie o hodnote 27V s maximálnym prúdom 1A. Je elektronicky chránený (bez poistiek) pred skratovaním obvodu a preťaženiu, zaplňa 8 modulov lišty DIN v rozvodnej skrini [5].

*Jednoreléový akčný člen F411/1N.* Akčný člen s možnosťou zapojenia do rozvádzačov s DIN lištou. Toto zariadenie obsahuje obojsmerné relé a tlačidlo pre miestne ovládanie záťaže, aktívne iba v prípade nakonfigurovania jednotky. Zariadenie môže byť nainštalované v systémoch MY HOME a môže byť nakonfigurované virtuálne aj fyzicky. Nakonfigurovaním všetkých pozícií môže byť priradený maximálne do troch oddelených skupín [5].

*Dvojreléový akčný člen F411/2.* Akčný člen pre inštaláciu na DIN lištu v rozvádzači. Toto zariadenie obsahuje dva nezávislé relé pre nezávislé spúšťanie záťaží, zahŕňa tiež lokálne ovládanie v podobe dvoch tlačidiel pre každý výstup zvlášť. Tieto tlačidlá sú aktívne len v prípade, že je jednotka úspešne nakonfigurovaná. Zariadenie sa inštaluje v systéme MY HOME a môže byť konfigurované virtuálne aj fyzicky. Dá sa použiť pre pripojenie motora koľajových záclon alebo roliet [5].

*Stmievač F414.* Akčný člen F414 napája odporové záťaže a feromagnetické transformátory. Po správnom zapojení stmievača na zbernicu a záťaž, môžeme ihneď ovládať intenzitu svetla z akéhokoľvek ovládacieho prvku, ak je správne nakonfigurovaný. Pre okamžité zapnutie a vypnutie svetla, stlačíme tlačidlo ovládača rýchlo, pre prispôsobenie intenzity svetla, podržíme tlačidlo ovládača dlhšie. Táto jednotka dokáže signalizovať rôzne problémy záťaže, ako napríklad rozbitie žiarovky a pod. Taktiež je chránená poistkou, ktorú v prípade vypálenia jednoducho vymeníme. Zariadenie môžeme nainštalovať ako súčasť MY HOME systému a môže byť konfigurované fyzicky aj virtuálne [5].

*Modul scenárov F420.* Tento prístroj umožňuje spravovať scenáre pre automatizáciu, ozvučenie a systému kontrolu teploty, ktoré boli vytvorené, pozmenené alebo aktivované, použitím rôznych zariadení v systéme automatizácie. V tomto module môže byť uložených maximálne 16 scenárov, pričom si každý scenár pamätá maximálne 100 nastavení. Scenáre tiež ovládajú odomykanie dverí a svietenie záhradného osvetlenia pri použití domového vrátnika v jednorodinnom dome. Na prednej strane zariadenia nájdeme dve tlačidlá a dva LED indikátory. Prvé tlačidlo uzamyká a odomyká proces programovania scenárov, s cieľom vyhnúť sa neúmyselným zmenám v uložených nastaveniach. Príslušná LED indukuje stav zapisovania: Zelená odomyknuté programovanie, červená uzamknuté programovanie, oranžová dočasné zablokovanie programovania. Druhé tlačidlo (DEL) ruší všetky scenáre, LED pod ním naznačuje, že zrušenie bolo úspešné, alebo že prístroj je v procese učenia [5].

*Pamäťový modul F425.* Pamäťový modul je pripojený na zbernicový systém kde neustále, po každom vyslanom signáli, ukladá stavy aktorov ktoré ovládajú svetlá. Toto zariadenie je veľmi užitočné v prípade dlhodobého výpadku dodávky elektrickej energie, alebo krátkych výpadkov (minimálne 400 milisekúnd), pretože dokáže obnoviť stav akčných členov svietidiel ako náhle nabehne dodávka elektriny. Obnova stavu členov trvá približne 10 sekúnd. Jeden pamäťový modul pripojený na zbernicu postačuje na pokrytie jedného systému (osobitný systém s jedným napájaním). Pri pospájaní viacerých systémov pomocou správne nastaveného modulu F422 taktiež stačí len jeden pamäťový modul pre všetky systémy. Pamäťový modul použijeme len na riadne nastavených systémoch [5].

*Dotyková obrazovka HW4890.* Dotyková obrazovka je zariadenie, ktoré umožňuje ovládanie funkcií MY HOME systému a správu osvetlenia pomocou jednoduchého a intuitívneho rozhrania 3,5“ dotykového LCD displeja. Zariadenie môže byť použité na správu automatizácie, osvetlenia, vykurovania, ozvučenia, zabezpečovacieho systému, hospodárenia energie a funkcií scenárov. Z každej skupiny je dotykové zariadenie schopné ovládať 20 členov [5].

*Tepelný termostat s prepínačom.* Zariadenie slúži na prispôsobenie teploty izby v zime aj v lete, pričom je schopné regulovať teplotu lokálne s ohľadom na teplotu určenú z centrálnej jednotky. Zariadenie má otočný gombík, pre miestne prispôsobenie teploty v rozmedzí 3°C od teploty centrálnej jednotky. Má tiež protimrazovú funkciu ako i režim vypnutia a zapnutia [5].

*SCS zbernicový kábel L4669.* Tento kábel je navrhnutý pre inštalácie systémom BUS pre ovládanie automatizácie vykurovania a správu spotreby energie. Kábel je možné použiť na napájanie systému (27V DC) a prenos signálov. Pri použití správnych izolačných svoriek a krytov, môže byť kábel položený do inštaláčnej rúrky alebo lišty spolu so silovými káblami [5].

#### IV. ZÁVER

Článok popisuje inteligentné inštalácie, zaeľuje riadiace systémy použité pri výstavbe alebo rekonštrukcii budov. Vymedzuje hlavné rozdiely v architektúre zbernicovej a klasickej elektroinštalácie, pričom poukazuje na pozitíva združovania riadenia pomocou zbernice. Zhŕňa hlavné funkcie komponentov a ich zapojenie, kde popisuje topológiu zbernicových systémov.

Popisuje konkrétnu značku BTicino, ktorá sa zaoberá automatizáciou budov. Spracúva tému spôsobu komunikácie komponentov, určuje topológiu a konkrétne menuje parametre jednotlivých komponentov potrebných pre návrh inteligentnej inštalácie, ktorá sa zaoberá osvetlením, správou zásuviek, automatizáciou vonkajších roliet a správou vykurovania.

Na základe skúmania konkrétnych komponentov potrebných pre pochopenie fungovania inštalácie, sme mohli prejsť na návrh inštalácie, kde sme postupne vyberali akčné členy pre jednotlivé miestnosti, následne sme vytvorili štruktúru kabeláže pre jednotlivé obvody, napájania komponentov inštalácie a komunikácie medzi nimi. Pre dosiahnutie správnej komunikácie, náš systém v stručných bodoch oboznamuje s konfiguráciou aktorov a ich ovládačov.

Na základe skúseností s vypracovaním projektu môžeme vysloviť záver, že práca projektanta pri návrhu inteligentnej inštalácie s týmto systémom je veľmi intuitívna a nie veľmi komplikovaná. Z užívateľského hľadiska modulárnosť tohto systému umožňuje veľmi jednoduché dopĺňanie riadiacich komponentov bez značného stavebného zásahu do objektu.

Najdôležitejšou výhodou tohto systému pre užívateľa je jednoduchosť ovládania a nastavenia funkcií zariadení inteligentnej inštalácie. V tomto zmysle je tento systém pre koncového užívateľa dobrou a návratnou investíciou.

#### LITERATÚRA

- [1] GAŠPAROVSKÝ, D.: Elektrotechnické a telekomunikačné inštalácie. Bratislava: Verlag Dashofer, 2008. ISSN 137-0324
- [2] Performance of building automation system in itelligent building. [online]. [cit. 07-05-2014]. dostupné na internete <<http://www.docstoc.com/docs/13474722/A-STUDY-ON-APPLICATION-AND-PERFOMANCE-OF-BUILDING-AUTOMATION-SYSTEM-IN-INTELLIGENT-BUILDING>>
- [3] Klasifikácia a charakteristika riadiacich systémov pre budovy [online]. [cit. 07-05-2014]. dostupné na internete: <<http://www.enoviny.sk/?sekcia=3&uroven=2&obsah=12&area=107&cid=17162>>
- [4] Klasická versus inteligentní elektroinstalace. [online]. [cit.08-05-2014f]. dostupné na internete: <<http://elektro.tzb-info.cz/domovni-elektroinstalace/7842-klasicka-versus-inteligentni-elektroinstalace>>
- [5] MY HOME automation. Design and installation guide. [online]. [cit.11-05-2014] dostupné na internete <<http://www.pageflip.bticino.it/myhomegb/>>

# Elektrická požiarňa signalizácia obchodného priestoru

<sup>1</sup>Matej JUŠČÁK, <sup>2</sup>Stanislav ILENIN

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
univerzita v Košiciach, Slovenská republika

<sup>1</sup>matej.juscak@student.tuke.sk, <sup>2</sup>stanislav.ilenin@tuke.sk

**Abstrakt** — Tento článok je zameraný na problematiku požiarnej signalizácie obchodného priestoru. V úvodnej časti článok popisuje problematiku požiarnej bezpečnosti a sumarizuje podklady potrebné pre správny návrh vhodného riešenia. V ďalšej časti je oboznámenie sa s návrhom elektrickej požiarnej signalizácie. Návrh je riešený podľa platných noriem, vyhlášok a predpisov.

**Kľúčové slová** — elektrická požiarňa signalizácia, požiarň hlásič, ústredňa

## I. ÚVOD

Pri riešení projektov dnešných novostavieb ako sú rôzne obchodné centrá, resp. administratívne budovy je jednou zo základných otázok správne zabezpečenie budovy. Riešenie tejto otázky je potrebné z hľadiska veľkej koncentrácie ľudí v pomerne malom, prípadne viacpodlažnom priestore. Je veľmi dôležité, aby v takýchto priestoroch bola zabezpečená ochrana ľudských životov pred vzniknutým požiarom. Z tohto dôvodu sa do budov inštaluje systém elektrickej požiarnej signalizácie, ktorý má včas odhaliť prípadný požiar a túto informáciu následne podať ďalej príslušným záchranným jednotkám, resp. obsluhu v závislosti od typu prevádzky EPS. Účelom EPS je tiež rýchla evakuácia v prípade vzniku požiaru.

## II. ROZBOR PROBLEMATIKY ELEKTRICKEJ POŽIARNEJ SIGNALIZÁCIE

Elektrická požiarňa signalizácia (EPS) je zariadenie pozostávajúce z ústredne, hlásičov požiaru, zariadení signalizujúce požiar, zariadení na prenos požiarnej signalizácie a z napájacieho zariadenia. EPS musí jednoznačne identifikovať najmenej jeden chemický alebo fyzikálny jav, ktorý je vyvolaný v dôsledkom požiaru v chránenom priestore. Tiež musí signalizovať poplach buď akusticky alebo opticky v chránenom priestore, resp. v jeho okolí, a musí ovládať zariadenia, ktoré sú k nej pripojené. V stavbe, ktorá je chránená pomocou EPS majú byť hlásiče požiaru umiestnené v priestore, ktorý je potreba chrániť. Do tohto chráneného priestoru sa umiestňujú tlačidlové hlásiče požiaru a automatické hlásiče požiaru. Samotné tlačidlové hlásiče požiaru (vynímajúc automatické hlásiče požiaru) sa môžu umiestniť do priestoru, v ktorom je zabezpečená trvalá prítomnosť osôb, alebo sú permanentne vytvárané podmienky, ktoré pôsobia nepriaznivo na spoľahlivosť automatických hlásičov požiaru. Automatickými hlásičmi požiaru musí byť v stavbe vybavený každý úsek, ktorého plocha je väčšia ako 1 000 m<sup>2</sup>, tiež ním musí byť vybavená lôžková časť zdravotníckeho zariadenia, ubytovacia časť hotela a inej stavby, ktorá je svojou charakteristikou určená na ubytovanie viac ako 100 osôb a vnútorný zhromažďovací priestor s kapacitou viac ako 500 osôb [1].

Elektrická požiarňa signalizácia musí byť navrhnutá tak, aby:

- každý požiar, ktorý vznikne bol signalizovaný ešte v počiatočnom štádiu pomocou samočinných hlásičov požiaru,
- pre každé miesto požiarneho úseku bolo zabezpečené účinné stráženie (čo najrovnomernejšie),
- bola vylúčená možnosť zníženia ich prevádzkovej spoľahlivosti,
- bol vylúčený falošný poplach,
- bol zabezpečený jednoduchý prístup k jednotlivým hlásičom pri údržbe, opravách a pod. [2].

Všetky požiadavky, ktoré sú kladené na elektrickú požiaru signalizáciu a tiež na návrh EPS sú zahrnuté v jednotlivých normách a vyhláškach. Pri správnom návrhu EPS treba dávať pozor na aktuálnosť týchto noriem a vyhlášok. Dané požiadavky na EPS sú obsiahnuté vo vyhláške MVSR č. 726 z roku 2002, ktorá ustanovuje vlastnosti EPS, podmienky za akých môže byť prevádzkovaná a určuje zabezpečenie pravidelnej kontroly EPS.

Elektrická požiaru signalizácia v sebe zahŕňa nasledujúce časti [3]:

- hlásiče požiaru (manuálny alebo automatický),
- požiarne slučky,
- ústredňu EPS,
- signalizačné linky,
- ostatné pomocné zariadenia ako ovládacie jednotky, signalizačné zariadenia a pod.

Pri návrhu a realizácii systému EPS sa musí postupovať podľa nasledujúcich noriem a vyhlášok:

- STN 73 0802/Z1, 2013 - Požiaru bezpečnosť stavieb. Spoločné ustanovenia [4].
- STN 73 0875, 1991 - Požiaru bezpečnosť stavieb. Navrhovanie elektrickej požiarnej signalizácie.
- STN P CEN/TS 54-14 (92 0404), 2005 - Elektrická požiaru signalizácia. Časť 16: Ústredňa hlasovej signalizácie požiaru.
- STN 92 0203/O1, 2013 - Požiaru bezpečnosť stavieb. Trvalá dodávka elektrickej energie pri požiaru [5].
- Vyhláška MV SR č. 726/2002 Z. z., ktorou sa ustanovujú vlastnosti elektrickej požiarnej signalizácie, podmienky jej prevádzkovania a zabezpečenia jej pravidelnej kontroly
- Vyhláška MV SR č. 94/2004 Z. z., ktorou sa ustanovujú technické požiadavky na protipožiaru bezpečnosť pri výstavbe a užívaní stavieb
- Vyhláška MV SR č. 307/2007 Z. z., dopĺňa vyhlášku MV SR č. 94/2004 Z. z.,
- Vyhláška MV SR č. 225/2012 Z. z., dopĺňa vyhlášku MV SR č. 307/2007 Z. z.

### III. SYSTÉM EPS

Jedným z riešení problému požiarnej bezpečnosti chránenej budovy je elektrický adresovateľný analógový protipožiaru systém FIREXA, tvorený ústredňou MHU 110, interaktívnymi hlásičmi a ostatnými potrebnými prvkami, ktoré svojimi vlastnosťami postačujú na bezpečné zabezpečenie budovy.

#### A. Ústredňa MHU 110

Je to zariadenie elektrickej požiarnej signalizácie určené k vyhodnocovaniu stavu aktuálnej požiarnej situácie v danom chránenom objekte. Ústredňa tohto typu využíva k riadeniu vnútorných a vonkajších funkcií mikroprocesory firmy Motorola. Jeden z mikroprocesorov je hlavný systémový, ostatné sú určené na riadenie liniek s hlásičmi.

Elektronika ústredne MHU 110 je zabudovaná v plechovej skrinke s vyklápacím vekom. V hornej časti skrine je umiestnený ovládacie panel spolu so signalizačným alfanumerickým displejom, signalizačné diódy a ovládacie tlačidlá. Vo vnútri skrine sú uložené lišty s elektronikou a prvky pre povrchovú montáž. Obsluha ústredne sa vykonáva pomocou klávesnice na štyroch prístupových úrovniach (podľa normy EN 54-2) zabraňujúcim zásah nepovolaných osôb do systému. Ústredňu je možné dodať v rôznych jazykových verziách vrátane slovenskej verzie.

Tento typ ústredne obsahuje dve kruhové linky, na každú kruhovú linku je možné pripojiť 128 adresovateľných prvkov (hlásičov). Navyše obe kruhové linky je možné rozdeliť na dve jednoduché linky. Prvky sú k vedeniu pripojené paralelne a vedenie týchto liniek je možné vetviť. Číslo, resp. adresa hlásiča sa nastavuje adresovacím prípravkom MHY 535 v prípade, že nemá mechanické nastavenie adresy [6].

#### B. Požiarne hlásiče

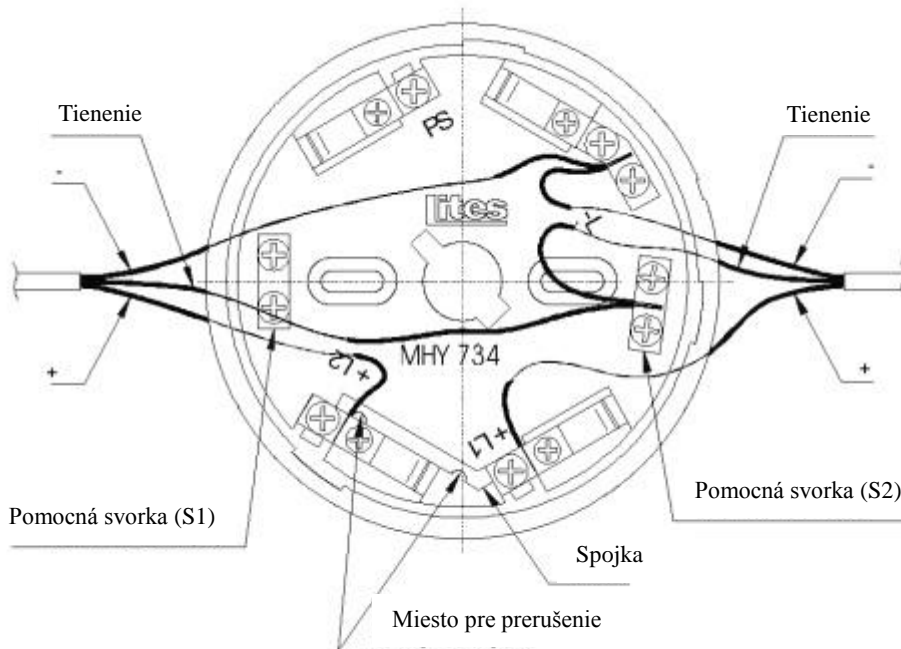
Optický hlásič dymu MHG 261 je interaktívny adresovateľný hlásič požiaru vyvinutý k spolupráci s ústredňami EPS LITES. Služi na signalizovanie vzniknutého požiaru ako detektor, ktorý reaguje na splodiny horenia na viditeľné aj neviditeľné častice dymu (aerosóly). Tento hlásič je určený na použitie vo vnútorných priestoroch chráneného objektu. V priestoroch kde sa hlásič použije sa nesmú vyskytovať agresívne látky.

Teplotný interaktívny hlásič teplôt MHG 362 sa využíva ako detektor, ktorý reaguje na nárast teploty alebo na dosiahnutie určitej preddefinovanej hodnoty teploty. Hlásič sleduje teplotu v danom prostredí a v jeho okolí, taktiež zmeny tejto teploty, aby na základe týchto informácií vyhodnotil situáciu.

Adresovateľný tlačidlový hlásič požiaru MHA 141 slúži na manuálnu signalizáciu požiaru osobou, ktorá požiar zistí. Inštaluje sa do priestorov schodísk, hál a všade tam, kde by automatické hlásiče požiaru boli neúčinné [6].

### C. Zásuvka MHY 734

Je určená na pripojenie vyššie spomenutých interaktívnych hlásičov požiaru. Je tvorená telesom z plastu, ktoré je osadené štyrmi perovými kontaktmi s príložkami na skrutky pre uchytenie vodiča. Tieto kontakty spájajú zásuvku s kontaktmi nožov hlásiča po jeho osadení. Zásuvka navyše obsahuje dva pomocné svorky pre možnú potrebu pripojenia ďalších vodičov.



Obr. 1 Zásuvka MHY 734 [6]

Svorky:

- -L prívod, vývod linky,
- +L1 prívod linky,
- +L2 vývod linky,
- PS paralelná signalizácia.

Navyše kontakty +L1 a +L2 sú prepojené spojku, ktorá vyraduje z funkcie izolátor v jednotlivých hlásičoch. V prípade potreby tohto izolátora sa prepojovacia spojka preruší pomocou klieští. Táto aktivácia izolátorov sa vykonáva až po oživení celého systému EPS [6].

### D. Tablo obsluhy MHS 811

Tablo obsluhy predstavuje prvok nadstavby interaktívneho protipožiarneho systému. Princípom činnosti je vyhodnocovanie informácií, ktoré sú získané z komunikačného kanála. Dosiahnuté informácie sa zobrazujú formou textu na displeji. Pracuje v dvoch režimoch, buď aktívne alebo pasívne. Pri nastavenom aktívnom režime dovoľuje tablo úplné ovládanie ústredne, zatiaľ čo pri pasívnom režime tablo len zobrazuje aktuálny stav ústredne, nemôže ju ovládať [6].

### E. Silový kábel N2XCH FE180

Bezhalogénový kábel N2XCH FE 180 so zvýšenou odolnosťou voči požiaru je určený pre pevný rozvod elektrickej energie. Využíva sa v miestach kde je potreba chrániť ľudské životy prípadne drahé prístroje v rôznych centrách, nemocniciach, letiskách a priemyslových komplexoch.



Jadro tohto kábla je zložené z jedného alebo viacerých holých medených drôtov a prídavnej izolácie zo sklo-sľudovej pásky: Izoláciu žil tvorí zosieťovaný polyetylén pričom žily sú stočené do duše kábla. Vonkajší plášť je tvorený ohňovzdorným polymérom oranžovej farby. Menovité napätie kábla je 0,6/1 kV, skúšobné napätie je 4 kV/50Hz a maximálny teplotný rozsah kábla pri skrate je 250°C po dobu 5 sekúnd.

#### IV. NÁVRH EPS

Zvolený systém pozostáva z automatických a manuálnych hlásičov požiaru. Väčšina priestorov z hľadiska chránenia bude opatrená automatickými hlásičmi požiaru (teplotnými, optickými). Manuálnymi hlásičmi (tlačidlové hlásiče) sa vybavujú priestory ako sú schodištia, chodby, možné cesty úniku a navyše priestory, v ktorých nie je vhodné použiť automatické hlásiče požiaru. Tieto tlačidlové hlásiče slúžia na ručné vyvolanie požiarneho poplachu. V EPS ústrední sa zhromažďujú informácie z detekčných zariadení, kde následným vyhodnotením týchto informácií bude spustený (resp. nespustený) požiarne poplach. Preto daný systém EPS musí obsahovať všetky potrebné zariadenia, ktoré sú potrebné na vyhodnotenie, indikovanie a signalizovanie požiarneho poplachu.

Chránený objekt sa rozdelí na 14 detekčných zón, podľa požiadaviek uvedených v norme STN P CEN/TS 54-14. Navyše z dôvodu zníženia porúch pri skrate musia byť medzi jednotlivé zóny v automatických hlásičoch požiaru umiestnené izolátory. Posledná štrnásť zóna bude tvorená všetkými manuálnymi hlásičmi požiaru.

Chránený objekt sa nerozdeľuje na žiadne zóny nakoľko v celej budove je poskytovaný len jeden poplachový (signalizačný) signál. Aby vzniknutá porucha na jednej linke neohrozila prevádzku všetkých signalizačných prvkov na dvoch podlažiach budú signalizačné prvky napájané samostatnými vstupmi.

Pri jednostupňovej signalizácii daná ústredňa hlási poplach hneď po popude z požiarneho hlásiča. Pri dvojstupňovej signalizácií bude vyhlásenie poplachu oneskorené o nastavené časy T1 a T2, z dôvodu aby obsluha zariadenia mohla správne overiť skutočný stav poplachu. Vyhodnotenie stavu prípadného poplachu je zabezpečené pracovníkom obsluhy tohto zariadenia, resp. preškoleným pracovníkom SBS služby, ktorí sa zdržiavajú v miestnosti v priestore tabla obsluhy.

#### V. ZÁVER

Popísaný protipožiarne systém s ústredňou MHU110, ktorá je zapojená v obvode spolu s ostatnými protipožiarne prvkami bezpečne chráni zvolený objekt. Pre správne pochopenie, návrh a samotné zapojenie tohto systému je potrebné sa riadiť inštrukciami a manuálom výrobcu EPS. Kombináciou týchto poznatkov a získaných poznatkov zo štúdie danej problematiky je možné vypracovať projekčný návrh s technickou dokumentáciou pre EPS.

#### LITERATÚRA

- [1] Vyhláška č. 726/2002 Z.z. Ministerstva vnútra Slovenskej republiky z 13. decembra 2002, ktorou sa ustanovujú vlastnosti elektrickej požiarnej signalizácie, podmienky jej prevádzkovania a zabezpečenia jej pravidelnej kontroly.
- [2] STN 73 0875: 1991, Navrhovanie elektrickej požiarnej signalizácie.
- [3] STN P CEN/TS 54-14: 2005, Elektrická požiarne signalizácia. Časť 14: Pokyny na plánovanie, projektovanie, inštalovanie, uvedenie do prevádzky, prevádzkovanie a údržbu.
- [4] STN 73 0802: 2010, Požiarne bezpečnosť stavieb.
- [5] STN 92 0203: 2013, Požiarne bezpečnosť stavieb. Trvalá dodávka elektrickej energie pri požiaru.
- [6] Lites a.s.: Elektrická požiarne signalizácia [online]. [cit. 2014-04-25]. Dostupné na internete: <<http://www.adiglobal.cz/>>

# Design of software tool for energy certification of buildings

<sup>1</sup>Július KOVÁCS, <sup>2</sup>Stanislav ILENIN

<sup>1</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>2</sup> Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>matagyula@gmail.com, <sup>2</sup>stanislav.ilenin@tuke.sk

**Abstract** — The goal of the paper is to create a software solution to help certify the lighting system for a range of buildings. At first a detailed look was taken at the technical background of the certification process, the legislative background and requirements. After understanding the different procedures and needs the software was developed using the Microsoft Visual Studio 2013 integrated development environment and the Visual Basic .NET programming language. The correctness and performance of the software solution was tested and demonstrated on an office building. The results were identical to or within the margin of error of those obtained from other sources.

**Keywords** — certification, energy performance, lighting

## I. INTRODUCTION

Energy performance certification provides a means of rating individual buildings, whether residential, commercial or public, on how (in)efficient they are in relation to the amount of energy needed to provide the residents with expected degrees of comfort and functionality. The degree of efficiency depends on many factors including: local climate, the design of the building, construction methods and materials, systems installed to provide heating, ventilation, air condition or hot sanitary water, and the appliances and equipment needed to support the functions of the building and its users. As it is apparent from all the aspects needed to be accounted for, certification is a complex procedure. It requires in-depth knowledge of building components. It also reflects increasing recognition of the need to think of building as integrated systems, rather than simply the sum of their parts. Energy certification is mandatory when selling a building, leasing a building or when a new building is completed or an old one renovated; otherwise it is voluntary. [1]

## II. ENERGY PERFORMANCE CERTIFICATION

The basis for energy certification is the calculation of energy efficiency and classification of buildings according to the so called categorization of buildings. Buildings are classified into energy efficiency classes from A to G. Buildings belonging to class A are the most energy efficient, while at the other end of the scale are the least energy efficient buildings in class G (they have high energy needs).

Many countries in Europe have increased the scale to include A1, A2 or A+, A++ or A- since the previous “A” boundary was found to be too limiting to reflect very efficient new buildings. Newly built buildings must meet the minimum energy efficiency requirements for new buildings, i.e. they must be classified in category A or B. If it is technically, functionally and economically feasible, an old building must comply with the minimum energy performance requirements for new buildings after a heavy renovation.

In general, two types of rating are used for building certification: asset rating is based on data derived from building inspection or drawings and building specifications; an operational rating uses metered data of actual energy consumption. Asset ratings are seen to be the most appropriate for new buildings and buildings in which there is frequent change of users, as the rating is independent of users and can be assessed before occupation.

An operational rating is more effective for buildings that have less frequent user turnover, and for large and complex buildings.

Energy certificates are valid for a maximum of ten years, unless there was a significant reconstruction of the building during that time. Classified as significant reconstruction of the building are construction modifications of the building which interfere with the buildings coating structure of more than 25% of its area, mainly facade or roof insulation and replacement of the original doors and windows. Energy certificates issued to 31.12.2012 are valid until the expiry there indicated.

Energy certification may be carried out only by a professionally qualified person for energy certification in accordance with law - the certificate of proficiency is issued by the Slovak Chamber of Civil Engineers (Slovenská komora stavebných inžinierov - SKSI).

The result of the energy performance certification is the energy performance certificate. Energy certificates display the calculated outputs of the energy assessment, thereby providing key information for all stakeholders for a given building. Certificates need a simple, straightforward layout to ensure clarity, ease of use and comparability for all users. The certification should nonetheless provide sufficient detail from the energy assessment and appropriate information upgrading for owners and building managers. These certificates are valuable to all stakeholders in the building sector. They provide a mechanism by which prospective buyers and tenants can compare the energy efficiency of different buildings or the energy rating across a range of similar buildings.

Certification also compares existing buildings to recent building codes, providing a way to compare existing and new buildings. In this regard, certificates are often considered a valuable piece of information at the time consumers are making decisions on property purchases, or rentals for either new or existing buildings. But certificates can also be valuable to sellers and property owners: buyers/renters might be attracted by the opportunity to save on energy bills by purchasing or renting a more efficient building. Or, they may opt to purchase/rent a less expensive building, knowing in advance that it is less efficient but can be improved through upgrades identified on the certificate.

### III. LEGISLATIVE FRAMEWORK

On 16th December, 2002, the European Parliament and Council accepted directive number 2002/91/EC about the energy performance of buildings - Energy Performance of Buildings Directive, abbreviated EPBD. The fundamental objective of the directive is to ensure the improvement of the energy performance of buildings. According to the directive, the climatic and local conditions should be taken into account in the process of improving the energy performance of buildings.

On 8th November, 2005, the National Council of the Slovak Republic passed law number 555/2005 Z.z. about the energy performance of buildings [2], which was prepared by the Ministry of Construction and Regional growth of the Slovak Republic. It encompasses all of the requirements of directive number 2009/91/EC. This law is in effect as of 1st January, 2008.

The technical standards concerning lighting systems are:

- STN EN 15193:2008 (36 0460) Energy performance of buildings. Energy requirements for lighting. [3]
- STN EN 15217:2008 (73 0720) Energy performance of buildings. Methods of expressing energy performance and energy certification of buildings. [4]
- STN EN 12251:2008 (12 8003) Input data of the indoor environment for the design and assessment of energy performance of buildings - quality of air, thermal state of the environment, lighting and acoustics. [5]
- STN EN 15459:2008 (06 0004) Energy performance of buildings. Economic assessment of energy systems. [6]

The basic methodological document is standard STN EN 15193, but other relevant norms are used at the different stages of the evaluation process, such as STN EN 12464-1 and others.

Individual countries can develop national methodologies which build upon the norms, adding region specific requirements or other specifics, such as unifying procedural steps. Slovakia has developed a detailed national methodology. For lighting systems it originates from STN EN 15193, which it supplements concerning partial calculations. It makes an algorithm of the whole process, which progresses step by step from known data to unknown data.

#### IV. SOFTWARE SOLUTION

The presented software solution is to be used in the process of energy performance certification, more precisely at the stage of determining the energy performance of the lighting system. It accepts user input for the list of rooms, windows, lighting systems and their various aspects and then calculates the lighting energy performance rating of the given building.

The implementation uses a tabbed graphical user interface to make data input easier. Beside the final results the results of partial calculations can be checked upon on the different tabs.

From the users point of view, the application needed a clean and easy to use graphical user interface. As large amounts of data have to be inserted in the process of certification, a table view of the data was chosen, creating a similar look and feel of spreadsheet processing software. The different categories are divided into multiple tabs. This makes it easier to keep track of what has already been entered, and also makes it easier to add new components to the application in future development.

Every calculation process is coded as a sub procedure, for example the sub procedure calcFD() calculates the factor of daylight availability. Each sub procedure is commented in the source code. This approach helps future proof the application, because if there is a change in one of the formulas only the appropriate sub procedure has to be modified, or re-written.

Each of the tabs in the user interface has a DataGridView field. The "Rooms" tab allows the user to add and delete rows from the DataGridView table. The rest of the tables get a set number of rows based on how many rows the user input in the "Rooms" tab. DataTables are used in the tabs "Calculations", "Calculations 2", "Monthly FD", "Monthly Power consumption (W)", "Results". DataTables make it easy to work with lots of data, because the addressing scheme allows to use unique column names, instead of just numbers.

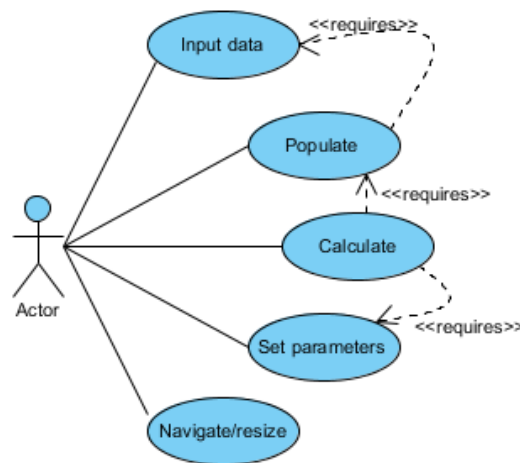


Fig. 1 Use case diagram.

The user can interact with the software as shown in the use case diagram in figure 1. The user can input data, set parameters, click the populate and calculate buttons, as well as navigate between the tabs and resize the application window. An example screenshot of the application can be seen in figure 2.

The screenshot shows the LightCalc software interface. At the top, there are input fields for Latitude (48, 44, 58, 48.75), Longitude (21, 15, 47, 21.26), and building name (B1 Office building). There are also fields for loss coefficients (0.7, 0.8, 0.85) and a 'Populate' button. Below this is a menu bar with 'Rooms', 'Windows', 'Lighting', 'Calculations', 'Calculations2', 'Monthly FD', 'Monthly Power consumption (W)', and 'Results'. The main area contains a table with the following columns: Room type, Room Length (m), Room Width (m), Room Height (m), Em (lx), Maintenance Factor, Height of the work plane (m), Angle of a linear obstacle (°), Angle of a horizontal obstacle (°), Angle of a vertical obstacle (°), and An (m<sup>2</sup>). The table lists various rooms such as Entry hall, Corridor, Storage, Toilets, Bathroom, Office for 2-6 persons, and Bigoffice2, with their respective dimensions and calculated values.

Room type	Room Length (m)	Room Width (m)	Room Height (m)	Em (lx)	Maintenance Factor	Height of the work plane (m)	Angle of a linear obstacle (°)	Angle of a horizontal obstacle (°)	Angle of a vertical obstacle (°)	An (m <sup>2</sup> )
Entry hall	1.55	2.20	2.79	100	0.8	0.85		37		3.4100
Corridor	9.83	3.55	2.79	100	0.8	0.85				34.8965
Storage	3.97	0.90	2.79	200	0.8	0.85				3.5730
Corridor	19.96	1.35	2.79	100	0.8	0.85				26.9460
Toilets	1.90	1.30	2.79	200	0.8	0.85				2.4700
Toilets	3.08	1.25	2.79	200	0.8	0.85				3.8500
Bathroom	5.52	2.90	2.79	200	0.8	0.85				16.0080
Office for 2-6 persons	5.56	2.90	2.79	500	0.8	0.85				16.1240
One table office	6.39	2.90	2.79	500	0.8	0.85				18.5310
Office for 2-6 persons	4.40	4.37	2.79	500	0.8	0.85				19.2280
One table office	4.93	3.65	2.79	500	0.8	0.85				17.9945
Office for 2-6 persons	13.20	3.65	2.79	500	0.8	0.85				48.1800
Corridor	1.50	7.13	2.79	100	0.8	0.85				10.6950
Office for 2-6 persons	5.15	4.19	2.79	500	0.8	0.85			66	21.5785
Storage	6.54	2.71	2.79	200	0.8	0.85				17.7234
Bigoffice2	28.79	11.88	3.30	500	0.8	0.85			73	342.0252

Fig. 2 LightCalc main window.

The function of the populate button is to prepare the DataTables necessary for further calculations based on the list of rooms input by the user - it adds the necessary number of new rows, as well as inserts the list of rooms into the respective tables for easier navigation. It also performs some of the necessary partial calculations for the "Calculations" tab.

The user is unable to interact with the populate and calculate buttons without inputting the necessary data (some rooms and their dimensions), this dependency is shown in the use case diagram with the <<requires>> tags. This is a design decision, as the populate button prepares the DataTables necessary for the calculations based on the list of rooms input by the user. If the user doesn't input any rooms, the DataTables can't be prepared, and without the DataTables the calculations will not be performed.

## V. PRACTICAL DEMONSTRATION

To demonstrate the application and to test the correctness of the calculated data, an office building was used. The building was the office building of ABB company located in Košice on Magnezitárska Street number 11. It is a two storey building and it has undergone reconstruction in 2002, thermally insulating the building and fitting plastic windows and doors all around. The cellar was not included in the energy certification process. The building can be seen in figure 3.



Fig. 3 ABB office building.

The list of rooms with all the necessary data, such as geometrical dimensions, type and count of lamps used, geological location, etc., were kindly provided by Mr. Radoslav Bumbal. Using the provided data the following results were obtained:

- the total annual power consumption of lighting for the whole building: 26236.05 kWh/year
- the power indicator for lighting (LENI): 17,78 kWh/(m<sup>2</sup>.year)
- the specific annual power consumption of lighting ( $\eta_E$ ): 0,04 kWh/m<sup>2</sup>.year.lx
- the final energy class for lighting: B.

A screenshot of the results can be seen in figure 4.

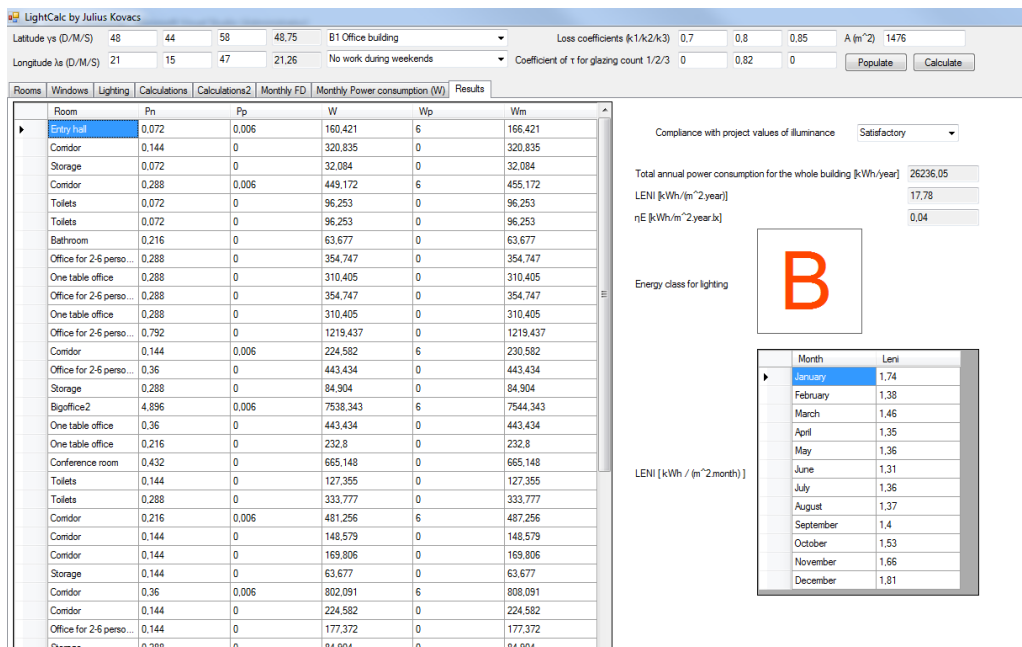


Fig. 4 The results tab

## VI. CONCLUSION

The application was able to successfully calculate the energy performance rating for lighting of an office building. One area where this application beats its spreadsheet counterpart is the hardware and software requirements. To work with the Microsoft Office version the user has to own and install the Microsoft Office software suite, which takes up hundreds of megabytes of space. The presented application is only 119 KB in size, even though the size of the redistributable .NET package adds another 90 MB, it is still smaller in size than the Microsoft Office software suite.

Despite this, the usefulness of application as a stand-alone tool has to be questioned, because just by itself it is not enough for a complete energy performance certification of a building. Other areas, such as heating and water supply have to be evaluated, too, and that is why it would server its purpose better if it were a part of a bigger software suite specialized on energy performance certification.

## REFERENCES

- [1] STERNOVÁ, Zuzana, a kolektív: Energetická hospodárnosť a energetická certifikácia budov. Bratislava, 2010, ISBN 978-80-8076-060-1.
- [2] Law number 555/2005 Z.z. about the energy performance of buildings.
- [3] STN EN 15193 Energy performance of buildings. Energy requirements for lighting, 2008 (36 0460).
- [4] STN EN 15217 Energy performance of buildings. Methods of expressing energy performance and energy certification if buildings, 2008 (73 0720).
- [5] STN EN 12251 Input data of the indoor environment for the design and assessment of energy performance of buildings - quality of air, thermal state of the environment, lighting and acoustics,2008 (12 8003).
- [6] STN EN 15459 Energy performance of buildings. Economic assessment of energy systems, 2008 (06 0004).

# Prenos informácií v elektrických staniaciach

<sup>1</sup>Maroš KACŮR, <sup>2</sup>Stanislav ILENIN

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>maros.kacur@student.tuke.sk, <sup>2</sup>stanislav.ilenin@tuke.sk

**Abstrakt** — Zameraním tohto článku je priblížiť detailnejšie spôsob komunikácie v elektrických staniaciach a časť komunikácie medzi elektrickými stanicami navzájom v minulosti a dnes. Poukazuje na nutnú obsluhu elektrickej stanice človekom v minulosti oproti dnešným automatizovaným bez potrebnej obsluhy. Znázorňuje základnú komunikačnú štruktúru automatizovanej elektrickej stanice a zariadení potrebných pre jej plnohodnotnú prevádzku. Bližšie špecifikuje jednotlivé komunikačné protokoly používané v energetickom priemysle, s podrobnejším opisom protokolu IEC 61850. V závere sa hlbšie venuje horizontálnej komunikácii v rámci stanice a simuluje základné prenosy GOOSE správ medzi ochranami navzájom s ich detailnejším sledovaním pomocou programu Wireshark.

**Kľúčové slová** — komunikácia, elektrická stanica, komunikačné protokoly, GOOSE správy

## I. ÚVOD

Klasické metódy prenosov dát a komunikácie nahrádzajú novšie spôsoby ako optické prepojenia, ktoré sú rýchlejšie a spoľahlivejšie. Komunikačné protokoly vytvorené v minulosti nahrádzajú nové, ktoré sú mnohonásobne rýchlejšie a ktoré spĺňajú podmienku vzájomnej komunikácie zariadení od rôznych výrobcov. Čoraz viac sa implementuje automatizovaná technika do elektrických staníc pre širšiu možnosť ovládania a sledovania prevádzkových stavov na diaľku.

V prvej časti tohto článku je opísaná história komunikácie a ovládanie jednotlivých prvkov elektrickej stanice v minulosti a dnes spolu s možnosťami komunikácie mimo elektrickej stanice. Sú definované jednotlivé typy komunikácie.

V druhej časti článku je opísaný najnovší medzinárodne používaný komunikačný protokol IEC 61850. Praktická časť simuluje vysielanie GOOSE správ medzi ochranami. V programe Schneider Electric MICOM S1 sa konfigurujú parametre týchto ochrán. Zachytávanie a sledovanie posielaných GOOSE správ medzi ochranami sa sleduje v programe Wireshark.

V závere sú zhodnotené prínosy nového komunikačného protokolu IEC 61850 a jeho využitie v prevádzke.

## II. HISTÓRIA KOMUNIKÁCIE

Komunikáciu v rámci elektrickej stanice a medzi elektrickými stanicami môžeme rozdeliť do troch historických etáp, v ktorých sa od manuálnej miestnej obsluhy postupne využívali čiastočné automatizované systémy až po dnešné plne automatizované systémy.

**1.stupeň:** ovládanie a manipulácie na mieste, metalické prepojenia na miestne relé, ručná manipulácia pomocou priamych elektrických impulzov, komunikačný kanál s dispečerským centrom len cez telefón.

**2.stupeň:** RTU (Remote Terminal Unit) zariadenia schopné zachytiť signál prenášaný elektrickým vedením a následne vykonať potrebný úkon (zap/vyp).

**3.stupeň:** nástup PC a elektronických zariadení, nové komunikačné protokoly (60870-5-101/104,61850), RTU zariadenia s ethernet portmi [1].

## III. KOMUNIKAČNÉ PROTOKOLY

Komunikačné protokoly sa v elektrických staniaciach používajú pre monitorovanie a ovládanie zariadení a samotnú komunikáciu medzi zariadeniami.

Komunikačné protokoly sa odlišujú svojimi vlastnosťami či už je to prenosová rýchlosť alebo metóda komunikácie t. j., či zariadenia komunikujú navzájom alebo zariadenie vysiela informácie nadradenému zariadeniu.

**Modbus** je otvorený komunikačný protokol, ktorý využíva komunikáciu typu master – slave. V praxi sa tento protokol používa pre komunikáciu medzi senzormi a nadradeným zariadením kde nadradené zariadenie vždy zahajuje komunikáciu. Tento protokol nie je vhodný pre diaľkové meranie a ovládanie [2].

Protokol **60870-5-101** je prevádzkovaný v dvoch režimoch. Vyvážený režim- kde sú stanice prepojené pomocou peer – to – peer komunikácie. Tento spôsob sa využíva hlavne v elektrických stanicach obsahujúce datakoncentrátor, ktoré následne preposielajú svoje dáta nadradenej stanici napr. dispečingu. Nevyvážený režim- stanice sú prepojené s nadradenou stanicou jednotlivito (využíva komunikáciu point – to - multipoint). Rozšírenou verziou tohto protokolu je **60870-5-104**, ktorý je rozšírený o možnosť komunikácie a ovládania jednotlivých zariadení pomocou IP adres prostredníctvom TCP/IP siete. Veľkou výhodou v praxi je možnosť komunikácie po verejnej sieti, avšak s nutnosťou šifrovania dát.

Najnovší medzinárodne používaný protokol je **IEC 61850** ktorý rýchlo našiel uplatnenie v energetike, kvôli spolupráci zariadení od rôznych výrobcov. Postupom času bol tento protokol rozšírený aj pre vodné elektrárne [3].

Tento protokol využíva metódu komunikácia klient – server, ale oproti klasickej komunikácie klient – server umožňuje aj klientskym zariadeniam riadiť prenos. Veľkým plusom tohto protokolu je peer – to – peer komunikácia, ktorá sa využíva v GOOSE alebo GSE správach. Každé zariadenie je reprezentované IP adresou, čím je umožnené sledovanie jeho prevádzkových stavov alebo jeho ovládanie. Zariadenia pre komunikáciu navzájom využívajú vysokorýchlostné GOOSE správy, ktoré sa zdefinujú pomocou IP adres iba pre tie zariadenia, ktoré majú tieto správy prijímať [4].

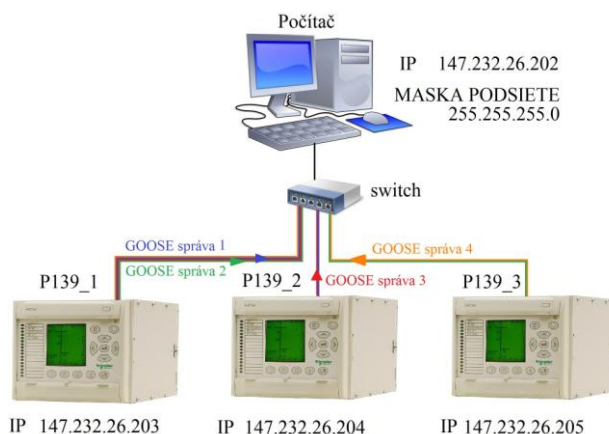
GOOSE správy sa delia do ôsmich typov, kde každý typ má presne zadefinovaný čas prenosu. Prvé dva typy definujú rýchle správy, ktoré sa líšia podľa toho, či sa jedná o prenosovú alebo distribučnú sústavu. Pre časovo kritickú horizontálnu komunikáciu sú určené GSE správy, ktoré sa delia na:

- GSSE – správy, ktoré predávajú informácie o zmenách stavu. Prenos a spracovanie týchto správ oproti GOOSE správam trvá dlhšie.
- GOOSE – dáta a hodnoty sú organizované do tzv. dáta setov, ktoré sú potom prenášané v jednotnom časovom intervale cez miestnu sieť LAN zariadeniam, ktoré sú nastavené pre prijímanie týchto správ. Na tieto správy sú kladené prísne požiadavky, pretože reakcia elektrických ochrán je závislá na rýchlych reakciách ostatných zariadení.

Každá GOOSE správa má svoju prioritu pre jej spoľahlivú a dostatočnú rýchlosť tzv. VLAN priority. Hodnoty sa nastavujú v rozmedzí od 1 po 7, kde 1 – 3 znamenajú nízku prioritu (časovo nekritické správy) a hodnoty v rozmedzí od 4 – 7 znamenajú vysokú prioritu (časovo kritické správy). Tieto funkcie sú predovšetkým používané vo switchoch, kde sa využívajú časovo kritické správy ako napríklad vypínacie signály.

#### IV. KOMUNIKÁCIA GOOSE SPRÁVAMI

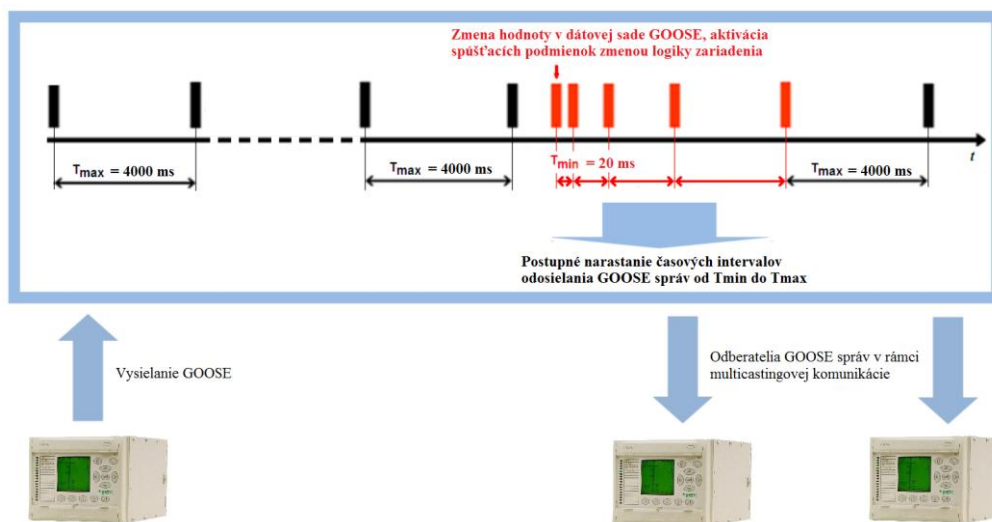
Na simulovanie komunikácie GOOSE správ sme použili tri ochrany značky MICOM prepojených navzájom prostredníctvom ethernetových káblov. Prijatie a odoslanie týchto správ bolo signalizované svietením a blikaním vopred prednastavených LED diód.



Obr. 1 Schéma zapojenia



Pre konfiguráciu nastavení jednotlivých ochrán bol použitý softvér Schneider Electric Micom S1 Studio. Nastavili sme IP adresy pre presné adresovanie každej správy. Ochrana s označením P139\_1 posielajú GOOSE správy 1 a 2 pre ochranu P139\_2 a P139\_3, ochrany P139\_2 a P139\_3 posielajú potvrdzujúce GOOSE správy 3 a 4 o doručení správ 1 a 2 pre ochranu P139\_1. Vyslanie správy je signalizované rozsvietením príslušných LED-iek a prijatie správy je signalizované blikaním LED-iek. Samotná simulácia pre zmenu, ktorá spôsobila vyslanie GOOSE správy, je zmena binárnej hodnoty. GOOSE správy sú posielané v časových cykloch, v našich ochránach bol tento čas 4000 ms. V okamihu poslania správy sa tento čas skrátil na 20 ms. Keďže prenášané dáta sa nemenia, tak sa tento čas začal postupne zvyšovať na 40 ms, 100 ms, 200 ms, 1000 ms, 2000 ms a nakoniec na pôvodnú hodnotu 4000 ms. Tieto časy, ktoré sa postupne zvyšovali sme sledovali v programe Wireshark. Grafické znázornenie tohto cyklického preposielania je na nasledujúcom obrázku. GOOSE správy majú podľa IEC 61850 preddefinovanú hodnotu VLAN priority 4, preto túto hodnotu v nastaveniach ani nemáme.



Obr. 2 Cyklické preposielanie GOOSE správ

## V. ZÁVER

V tomto článku sme prezentovali komunikáciu elektrických ochrán pomocou GOOSE správ a detailnejšie sme túto komunikáciu sledovali v programe Wireshark, kde sme si overili cyklické preposielanie dát v okamihu prenosu GOOSE správ a následné zvyšovanie tohto časového cyklu na pôvodnú hodnotu. V praxi sú tieto správy využívané v blokovacích podmienkach, oznámenie záložnej ochrane v prípade zlyhania hlavnej ochrany alebo informácia o zmene konfiguračných súborov, ktoré sú vopred nadefinované. Výhodou týchto správ je schopnosť prijímateľom zistiť náhle prerušenie komunikácie alebo preťaženie komunikačného spojenia prípadne strate dát, čo je v praxi veľmi dôležité.

## LITERATÚRA

- [1] ČECH, Marek: *Komunikácia elektrických ochrán*, Diplomová práca, FEI TU Košice, 2010, 96s.
- [2] RONEŠOVÁ, Andrea: *Přehled protokolu MODBUS*, 2005. Dostupné na internete: <<http://home.zcu.cz/~ronesova/bastl/files/modbus.pdf>>
- [3] IEC 61850 Part 1 – Part 10:2002-2014, *Communication networks and systems in substations*
- [4] COBUS, Strauss.: *Electrical Network Automation and Communication Systems*, 2003, 201 s. ISBN 978-0-7506-5801-0. Dostupné na internete: <<http://www.scribd.com/doc/8545096/Practical-Electrical-Network-Automation-and-Communication-Systems>>

# Diskriminačná analýza reči založená na adaptívnom modelovaní foném

*Peter VISZLAY, Michal HUDÁK, Jozef JUHÁR*

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky, Technická univerzita v Košiciach, Slovenská republika

peter.viszlay@tuke.sk, michal.hudak12@gmail.com, jozef.juhar@tuke.sk

**Abstrakt** – Definícia tried v lineárnej diskriminačnej analýze (LDA) v systémoch automatického rozpoznávania reči (ARR), je stále aktuálnou témou. Nie v každej úlohe je zrejmé, akým spôsobom je potrebné definovať LDA triedy, aby bola požiadavka ich dostatočnej diskriminácie splnená. V tomto článku je prezentovaný nový spôsob definície tried v LDA, ktorý využíva vybrané fonetické vlastnosti rečových jednotiek - foném. Niektoré časové vlastnosti foném, ako je nábeh fonémy, hlavné a záverečné trvanie, sú využité na definíciu „hlavných“ a tzv. „prechodových“ tried. Hlavné triedy sú príznačné pre klasickú LDA, zatiaľčo prechodové triedy obsahujú úseky reči adaptívne zvolené pomocou našich prístupov. Tieto dve skupiny tried sú potom v LDA modelované spoločne ako aj oddelene, pričom je skúmaný vplyv na úspešnosť LDA. Bolo vytvorených niekoľko testovacích scenárov s odlišnou konfiguráciou navrhnutých postupov. Tieto boli evaluované v úlohe automatického rozpoznávania plynulej reči s veľkým slovníkom pre slovenčinu na báze trifón s využitím slovenského rečového korpusu Parlament. Vytvorené akustické modely sú porovnané s referenčným modelom MFCC (Mel-frequency Cepstral Coefficients) a LDA trénovaného tradičným spôsobom. Výsledky experimentov deklarujú vylepšenie úspešnosti rozpoznávania obidvoch referenčných modelov.

**Príučové slová** – fonéma, lineárna diskriminačná analýza, modelovanie, rozpoznávanie, segmentácia

## I. ÚVOD A MOTIVÁCIA

Lineárna diskriminačná analýza (LDA) [1], [2], [3] je jednou z najznámejších lineárnych transformácií používaných v systémoch automatického rozpoznávania reči (ARR) na dodatočné vylepšenie diskriminačných vlastností základných príznakových vektorov reči. Dôležitou časťou LDA je definícia tried, pri ktorej sa musí špecifikovať, aké rečové jednotky s podobnými vlastnosťami sa priradia do rovnakej triedy. V praxi existuje mnoho navrhnutých prístupov k definícii tried, avšak nie každá je pre danú úlohu rozpoznávania vhodná. Definícia tried závisí od konkrétnej úlohy rozpoznávania, od náročnosti aplikácie a čiastočne aj od modelovaného jazyka a veľkosti slovníka ako aj veľkosti trénovacích rečových dát.

Medzi najpoužívanejšie rečové jednotky patria fonémy, ktoré je možné vhodne modelovať. V anglickom jazyku je známych približne 50 foném a v slovenskom jazyku približne 45 [4]. Fonémy sú v podstate nezávislé od úlohy a môžu byť trénované na jednu úlohu a testované na inú. Pri takomto type modelovania však nie sú do úvahy brané koartikulačné vlastnosti. Jeden zo spôsobov modelovania koartikulačných vlastností je použitie väčších rečových jednotiek (napr. trifón) [5]. Tieto jednotky zahŕňajú skupinu foném, ktoré obsahujú kontextuálne vlastnosti.

Vzhľadom k tomu, že prechody z fonémy na fonému sú zle modelované pomocou fonémových modelov, je jedným z riešení modelovanie týchto prechodov. Experimentálna práca prezentovaná v tomto článku je orientovaná práve na túto problematiku. Navrhujeme nový prístup k definícii tried v LDA pre slovenské fonémy založený na adaptívnom modelovaní foném využívajúci dynamické pridelenie príznakových vektorov do LDA tried, pričom sú zohľadnené prechodové vlastnosti susediacich foném. Je ukázané, že takýto prístup k modelovaniu vedie k vylepšeniu štandardných postupov v LDA. V práci sme boli čiastočne motivovaní našimi predchádzajúcimi prácami [6] a [7], v ktorých sme sa partikulárne venovali tejto oblasti.

Štruktúra článku je rozdelená nasledovným spôsobom. Druhá kapitola opisuje metódu LDA, tretia kapitola sa detailnejšie venuje opisu navrhovaného adaptívneho modelovania a štvrtá kapitola prezentuje opis systému ARR. Piata kapitola prezentuje a analyzuje experimentálne výsledky a v šiestej kapitole je uvedený záver.

## II. LINEÁRNA DISKRIMINAČNÁ ANALÝZA

V systémoch ARR sa na lineárnu transformáciu využíva mnoho techník. Medzi najznámejšie patrí analýza hlavných komponentov (v angl. „Principal Component Analysis“ - PCA) [8] a lineárna diskriminačná analýza (v angl. „Linear Discriminant Analysis“ - LDA) [1], [2]. V tejto práci je pozornosť venovaná metóde LDA, ktorá hľadá optimálnu lineárnu transformáciu  $G$ , pomocou ktorej sa transformujú originálne  $n$ -rozmerné dáta do priestoru s rozmerom  $p$ ,  $p < n$ . Predpokladajme, že originálne vstupné tréningové dáta v matici  $A$  sú rozčlenené do  $k$  tried  $A = \{\Pi_1, \dots, \Pi_k\}$ , kde trieda  $\Pi_i$  obsahuje  $n_i$  dátových bodov z  $i$ -tej triedy, pričom  $n = \sum_{i=1}^k n_i$ .

Potom optimálna transformácia  $G$  by mala zachovávať štruktúru tried originálneho vysoko-rozmerného priestoru v novom nízkorozmernom priestore. Cieľom metódy je nájsť takú lineárnu schému, ktorá maximalizuje *separáciu tried* v redukovanom priestore tým, že sa maximalizuje *medzitriedny rozptyl* a minimalizuje sa *vnútrotriedny rozptyl*, čím sa dosiahne maximálna diskriminácia [4], [6]. Na kvantifikovanie týchto rozptylov sa využívajú rozptylové matice. *Vnútrotriedny rozptyl* (v angl. „within-class scatter“) je určený vnútrotriednou rozptylovou maticou  $S_W$  nasledovne:

$$S_W = \sum_{i=1}^k \sum_{x \in \Pi_i} (x - \mu_i)(x - \mu_i)^T \quad (1)$$

a *medzitriedny rozptyl* (v angl. „between-class scatter“) je charakterizovaný medzitriednou rozptylovou maticou  $S_B$  nasledovne:

$$S_B = \sum_{i=1}^k n_i (\mu_i - \mu)(\mu_i - \mu)^T, \quad (2)$$

kde

$$\mu_i = \frac{1}{n_i} \sum_{x \in \Pi_i} x \quad (3)$$

predstavuje vektor strednej hodnoty pre triedu  $i$  (v angl. „class centroid“, „class mean“) a vektor

$$\mu = \frac{1}{n} \sum_{i=1}^k \sum_{x \in \Pi_i} x \quad (4)$$

predstavuje vektor globálnej strednej hodnoty (v angl. „global centroid“, „global mean“). Transformačná matica  $G$  je potom vypočítaná spektrálnym rozkladom matice  $S_W^{-1} S_B$  a dáta sa transformujú podľa vzťahu  $y = G^T x$ , kde  $x$  je originálny vektor a  $y$  je transformovaný vektor.

## III. ADAPTÍVNE MODELOVANIE FONÉM V LDA

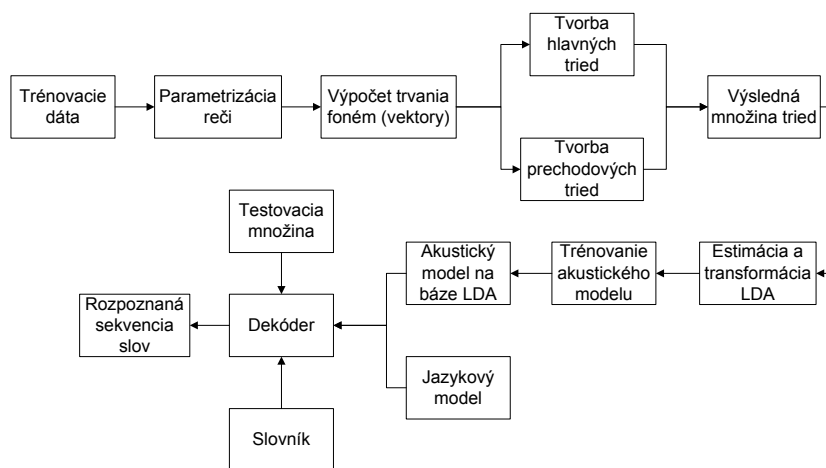
### A. Všeobecný opis

Proces adaptívneho modelovania foném je znázornený pomocou blokovej schémy na Obrázku 1. Hlavnú časť tvorí definícia tried, ktorú je možné charakterizovať blokmi pre tvorbu hlavných a prechodových tried. Definícia tried pre LDA má značný podiel na úspešnosti celého systému ARR. Najčastejším prípadom definície tried pre LDA je definícia, kedy triedy korešpondujú priamo s fonémami. Zo spektrografických vlastností hlások opísaných v [7] je známe, že trvanie významnej väčšiny foném sa skladá z troch častí (stavov) - *explózia* (nábeh fonémy), *plózia* (hlavné trvanie) a *implózia* (záverečné trvanie). Nie je možné so 100% presnosťou určiť hranice medzi fonémami, keďže prechody medzi fonémami (*expózia* - *implózia*) sa môžu prekrývať. Nie je možné tiež tvrdiť, že pri každom prechode medzi fonémami nastáva prekrývanie.

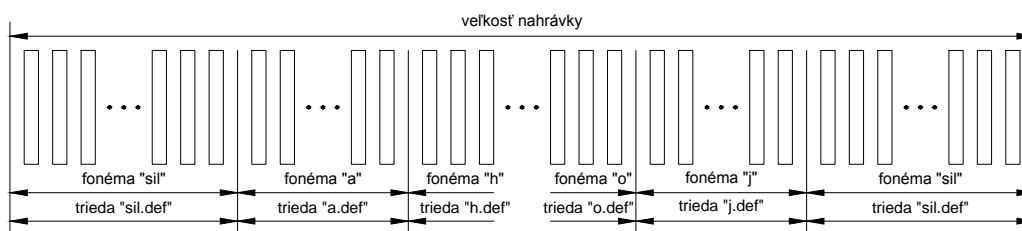
V dôsledku toho je náš návrh definície tried v LDA založený na *kombinácii hlavných (fonetických) tried s prechodovými triedami*. Pojem *hlavná trieda* znamená, že príznakové vektory prislúchajúce rovnakej fonéme sú sústredené do spoločnej triedy bez ohľadu na uvádzané časti trvania foném. U niektorých verzí môžu byť hlavné triedy reprezentované kombináciou stavov *plózia* - *implózia*, *explózia* - *plózia* alebo len jediným stavom a to *plóziou* (viď Obrázok 3). Tento obrázok znázorňuje distribúciu vektorov vo vyslovenej sekvencii *ahoj*, ktorá začína a končí úsekom ticha (*sil*).

### B. Tvorba fonetických tried

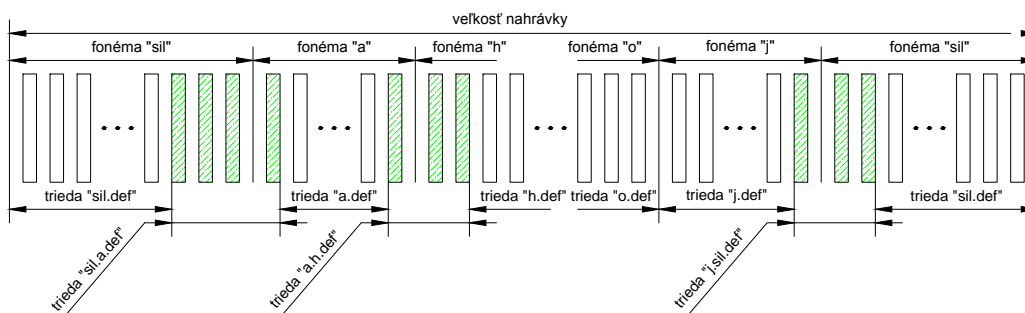
Vytváranie tried sa realizuje pridelením príznakových vektorov z daných nahrávok jednotlivým triedam na základe časového zarovnania foném pochádzajúceho z tréningového procesu. Princíp vytvárania *hlavných tried* spočíva v tom, že každej fonéme s rovnakým popisom (značkou) sa priradí jedna hlavná trieda s rovnakým popisom. To znamená, že príznakové vektory reprezentujúce danú fonému budú pridelené jej korešpondujúcej hlavnej triede. Definícia hlavných tried je graficky znázornená na Obrázku 2.



Obr. 1 Bloková schéma systému ARR využívajúceho adaptívne modelovanie slovenských foném v LDA.



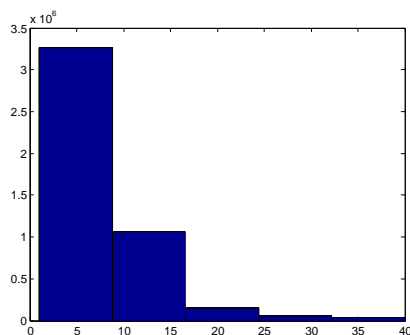
Obr. 2 Princíp definície hlavných tried vo vyslovenej sekvencii ahój.



Obr. 3 Princíp definície hlavných a prechodových tried s rôznymi podmienkami vytvárania tried v sekvencii ahój.

Naopak, pri vytváraní *prechodových tried*, daná trieda reprezentuje prechod medzi dvojicou za sebou nasledujúcich foném. Počet pridelovaných vektorov môže byť rôzny, v závislosti od veľkosti danej fonémy, z ktorej sú pridelované príznakové vektory. Podrobnejší opis pridelovania počtu vektorov triedam na základe zvolených intervalov (viď Tabuľka 1), je opísaný v Kapitole III-C. Grafické znázornenie definície prechodových tried je uvedené na Obrázku 3.

Na základe zvolených podmienok môžu v určitých trénovaných modeloch nastať situácie, kedy hlavná trieda bude pozostávať z príznakových vektorov, ktoré neboli pridelené prechodovej triede. Naopak pri prechodových triedach môže nastať prípad, kedy pre dvojicu za sebou nasledujúcich foném sa prechodová trieda nebude vytvárať z dôvodu krátkeho trvania fonémy. Takýto spôsob definície tried je znázornený na Obrázku 3. Pod pojmom krátka fonéma sa rozumie fonéma, ktorej trvanie je reprezentované tromi a menej príznakovými vektormi (v niektorých modeloch dvoma a menej príznakovými vektormi). Pri fonémach s takou krátkou dĺžkou nie je prakticky možné rozčleniť fonému na tri základné časti (explózia, plózia, implózia) s dostatočným trvaním. Z dôvodu analýzy vplyvu tohto faktu na úspešnosť rozpoznávania bolo snahou natrénovať viacero modelov, ktoré zahŕňali rôzne spomínané vlastnosti/podmienky pridelovania a vytvárania hlavných alebo prechodových tried.



Obr. 4 Štatistické rozloženie početnosti foném príslušných dĺžok v celej databáze pomocou histogramu.

Tabuľka 1

Počet vektorov  $m$  pridelovaných triedam na základe dĺžky fonémy  $M$  pre stochastický a histogramový princíp.

Stochastický rozsah vektorov	Počet pridelovaných vektorov	Rozsah vektorov na základe histogramu
$1 \leq M \leq 3$	$m = 0$	$1 \leq M \leq 3$
$4 \leq M \leq 7$	$m = 1$	$4 \leq M \leq 8$
$8 \leq M \leq 14$	$m = 2$	$9 \leq M \leq 16$
$15 \leq M \leq 29$	$m = 3$	$17 \leq M \leq 24$
$30 \leq M \leq 49$	$m = 4$	$25 \leq M \leq 32$
$M \geq 50$	$m = 5$	$M \geq 33$

### C. Histogramová analýza

Počet pridelovaných príznakových vektorov prechodovým triedam je rôzny a závisí od dĺžky aktuálnej fonémy. Na základe toho bolo potrebné určiť intervaly s príslušným počtom pridelovaných vektorov, ktoré môžeme vidieť v Tabuľke 1. Určovanie týchto intervalov má experimentálny charakter. V prvom prípade bolo určenie týchto intervalov na základe predpokladov z práce [7]. V druhom prípade bolo na určenie intervalov využité štatistické rozloženie pravdepodobností početnosti foném pomocou histogramu (viď. Obrázok 4). Oba prípady boli podložené štatistickými údajmi o výskyte početností foném v databáze. Tieto štatistické údaje obsahujú početnosť výskytu každej fonémy pre ich všetky možné dĺžky vyskytujúce sa v databáze. Je možné si všimnúť, že stochastická voľba hraníc mierne koreluje so štatisticky podloženou.

## IV. NASTAVENIE SYSTÉMU ARR A PODMIENKY EXPERIMENTOV

### A. Rečový korpus a predspracovanie reči

Na evaluáciu implementovaných postupov bola zvolená slovenská rečová databáza *ParDat1* [9], obsahujúca približne 100h spontánnych parlamentných debát. Databáza bola nahraná od 120 rečníkov (približne 90% mužov) v Národnej rade Slovenskej republiky pomocou stolových mikrofónov. Dĺžka trvania nahrávok je v rozsahu od 10s až po niekoľko desiatok sekúnd. Tréningovú množinu tvorilo 36917 nahrávok a testovaciu 884 iných foneticky bohatých nahrávok.

Nahrávky boli nahraté s pôvodnou vzorkovacou frekvenciou 48kHz, ktorá bola neskôr nastavená na 16kHz podvzorkovaním. Všetky tréningové a testovacie nahrávky boli predspracované podľa nasledovného postupu. Na rečový digitalizovaný signál sa v prvom kroku aplikoval filter s preemfázovým koeficientom 0,97. V ďalšom kroku sa rečový signál segmentoval na časové mikrosegmenty so štandardným trvaním 25ms s využitím Hammingovej oknovej funkcie. Časový interval vytvárania rámcov bol nastavený na 10ms. Počas melovskej analýzy bolo v banke filtrov využitých 26 kanálov. Výsledkom tohto postupu boli 26-rozmerné FBANK (v angl. „Filter-bank“) vektory. Po aplikácii DCT transformácie bolo zanechaných prvých 12 kepstrálnych koeficientov a bol pridaný nultý kepstrálny koeficient. Pri referenčnom akustickom modelovaní boli tieto 13-rozmerné vektory ďalej rozšírené o delta a akceleračné koeficienty, výsledkom čoho boli 39-rozmerné MFCC vektory. Na spracovanie dát pomocou LDA metód boli využité len základné 13-rozmerné MFCC vektory. Po lineárnej transformácii bol ich rozmer zanechaný a počas akustického modelovania bol expandovaný na rozmer 39 pomocou delta a akceleračných koeficientov.

### B. Topológia HMM a akustické modelovanie

Systém rozpoznávania reči s veľkým slovníkom používal kontextovo závislé trifóny modelované pomocou trojstavových ľavo-pravých skrytých Markovových modelov. Počet Gaussových zmesí

Tabuľka 2  
 Vplyv histogramovej analýzy pri definícii tried na úspešnosť LDA pre obe množiny tried

Počet zmesí	Acc. all	Acc. all hist.	rozdiel v %	Acc. cross	Acc. cross hist.	rozdiel v %
1	76,08	76,11	<b>+0,03</b>	76,12	76,20	<b>+0,08</b>
2	77,72	77,27	-0,45	77,23	77,16	-0,07
4	78,99	79,44	<b>+0,45</b>	78,91	78,87	-0,04
8	80,70	81,11	<b>+0,41</b>	80,49	80,69	<b>+0,20</b>
16	82,19	82,65	<b>+0,46</b>	82,49	82,65	<b>+0,16</b>
32	83,60	84,04	<b>+0,44</b>	83,73	83,85	<b>+0,12</b>

použitých na 1 stav bol expandovaný mocninou dvoch; od 1 po 32. Počet trifónových modelov korešpondoval s počtom trifón. Pre testovacie účely bol vytvorený bigramový jazykový model. Slovník systému pozostával približne zo 125000 slov. Trifónové LDA triedy zodpovedali 1294 trifónam vybraných z celkového počtu 18868 vyskytujúcich sa v trénovacej množine. Selekcia trifón bola založená na mapovacej procedúre opísanej v [9] z dôvodu problematickej LDA a 2DLDA analýzy úplného počtu trifón. Rovnaká mapovacia schéma bola dodržaná pri trénovaní aj testovaní akustických modelov. Parametrizácia reči, ako aj akustické modelovanie, boli vykonané použitím sady nástrojov HTK (v angl. „*Hidden Markov Model Toolkit*“) [10].

### C. Evaluácia systému

Z dôvodu evaluácie navrhnutých postupov bol zvolený parameter presnosť, resp. úspešnosť rozpoznávania Acc. (v angl. „*accuracy*“) v percentách, ktorá bola vypočítaná nasledovne:

$$Acc. = \frac{H - I}{N} \times 100\%, \quad (5)$$

kde symbol  $H$  vyjadruje počet správne rozpoznaných slov,  $I$  je počet vložených chýb (v angl. „*insertions*“) a  $N$  vyjadruje počet všetkých referenčných slov [10].

## V. EXPERIMENTÁLNE VÝSLEDKY A VYHODNOTENIE

Výsledky experimentov sú rozdelené do dvoch častí. V prvej sú vyhodnotené výsledky najúspešnejšej metódy, ktorá je založená na histogramovej analýze a v druhom prípade je uvedené globálne porovnanie úspešnosti celej navrhovanej metódy.

### A. Histogramová analýza v procese definície tried

V tejto podkapitole sú uvedené výsledky demonštrujúce pozitívny vplyv histogramovej analýzy použitej na určenie segmentačných hraníc foném. Uvedené výsledky základného modelu (bez využitia histogramovej analýzy) prislúchajú definícii tried v LDA, kedy bolo vykonané dynamické pridelovanie príznakových vektorov do fonetických tried na základe vlastností aktuálnej fonémy.

V Tabuľke 2 sú uvedené výsledky úspešnosti rozpoznávania na báze LDA využívajúcej adaptívne modelovanie foném s dynamickým pridelovaním vektorov do hlavných aj prechodových tried („all“, stĺpec 2, 3 a 4) a s pridelovaním len do prechodových tried („cross“, stĺpec 5, 6 a 7), pričom stĺpce 2 a 5 predstavujú výsledok bez histogramovej analýzy a stĺpce 3 a 6 predstavujú výsledok s histogramovou analýzou. V prípade vytvárania všetkých tried došlo k vylepšeniu základných modelov takmer pre každý počet zmesí, až na jeden prípad, kedy nastalo mierne zhoršenie. Najvyššia hodnota zlepšenia oproti základnému modelu bola o 0,46%. V prípade vytvárania len prechodových tried došlo k zlepšeniu základnej metódy takmer pre všetky počty zmesí. Výnimkou však bol počet zmesí 2 a 3 kedy nastalo zhoršenie o -0,07% a -0,04%. Najlepšie zvýšenie úspešnosti v porovnaní so základným modelom bolo o 0,20%.

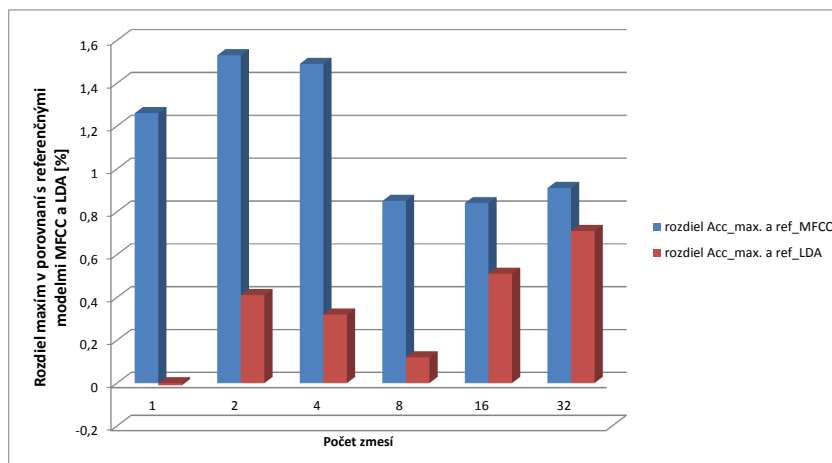
### B. Globálne porovnanie navrhovanej metódy s referenčnými modelmi

Nasledujúca podkapitola zahŕňa globálne porovnanie navrhovaného prístupu s referenčnými modelmi MFCC a LDA bez ohľadu na typ použitých tried v LDA alebo na využitie histogramovej analýzy. Napriek tomu, že akustické modely sú trifónové, definícia tried v referenčnej aj modifikovanej LDA je založená na fonémach. V Tabuľke 3 sú uvedené výsledné hodnoty referenčných modelov spolu s maximálnymi výslednými hodnotami z modelov založených na adaptívnom modelovaní foném. Posledné dva stĺpce tabuľky reprezentujú absolútne rozdielové hodnoty medzi výsledkami referenčných modelov a modifikovanej LDA. Najvýraznejšie vylepšenie úspešnosti rozpoznávania voči referenčnému modelu MFCC dosiahlo hodnotu +1,53% a v porovnaní s referenčným modelom LDA +0,71%. Porovnanie všetkých rozdielových hodnôt medzi sebou pre obidva referenčné modely je graficky znázornené na Obrázku 5.



Tabuľka 3  
 Porovnanie výsledkov navrhovaného princípu s referenčnými modelmi MFCC a LDA

Počet zmesí	Acc. MFCC	Acc. LDA	Acc. Max.	rozdiel MFCC	rozdiel LDA
1	74,94	76,21	76,20	<b>+1,26</b>	-0,01
2	76,19	77,31	77,72	<b>+1,53</b>	<b>+0,41</b>
4	77,95	79,12	79,44	<b>+1,49</b>	<b>+0,32</b>
8	80,26	80,99	81,11	<b>+0,85</b>	<b>+0,12</b>
16	81,81	82,14	82,65	<b>+0,84</b>	<b>+0,51</b>
32	83,11	83,31	84,04	<b>+0,93</b>	<b>+0,73</b>



Obr. 5 Grafické porovnanie absolútnych vylepšení referenčných modelov pomocou navrhovaného modelovania.

## VI. ZHRNUTIE A ZÁVER

Z výsledkov je možné vyvodit' záver, že navrhnuté adaptívne modelovanie foném s dynamickým pridel'ovaním vektorov na báze histogramovej analýzy prináša vylepšenie obidvoch referenčných modelov. Konkrétne, väčšie vylepšenie MFCC modelu nastáva pri nižších počtoch zmesí (1-4), zatiaľ čo vylepšenie konvenčného LDA modelu nastáva pri stredných počtoch zmesí (4-8). Tento výsledok zároveň potvrdzuje aj fakt, že v systémoch ARR sa so zvyšujúcim počtom zmesí obtiažnejšie dosahuje výrazné vylepšenie. V blízkej budúcnosti plánujeme detailnejšie sa venovať zahrnutiu ďalších vlastností foném do adaptívneho modelovania. Taktiež plánujeme využiť inú štatistickú analýzu (okrem histogramovej) na určenie segmetačných hraníc foném.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu UVP Technicom (kód ITMS projektu: 26220220-182) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] R. Haeb-Umbach and H. Ney, "Linear discriminant analysis for improved large vocabulary continuous speech recognition," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'92)*, San Francisco, CA, March 1992, pp. 13–16.
- [2] R. Haeb-Umbach, D. Geller, and H. Ney, "Improvements in connected digit recognition using linear discriminant analysis and mixture densities," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'93)*, Minneapolis, MN, April 1993, pp. 239–242.
- [3] J. Pyllkkönen, "LDA based feature estimation methods for LVCSR," in *Proc. of the International Conference on Spoken Language Processing, (INTERSPEECH'06)*, Pittsburgh, USA, September 2006, pp. 389–392.
- [4] J. Juhár *et al.*, *Rečové technológie v telekomunikačných a informačných systémoch*. Košice: EQUILIBRIA, s.r.o., 2011.
- [5] K.-F. Lee, *Automatic Speech Recognition: The Development of the Sphinx Recognition System*. Springer, 1989, vol. 62.
- [6] P. Vizslay, "Lineárne transformácie príznakového priestoru v systémoch automatického rozpoznávania reči," 2010, písomná práca k dizertacnej skúške.
- [7] M. Hudák, "Modelovanie spektrografických vlastností slovenských hlások v lineárnej diskriminačnej analýze," 2012, bakalárska práca.
- [8] I. T. Jolliffe, *Principal Component Analysis*. New York, USA: Springer-Verlag, 1986.
- [9] S. Darjaa *et al.*, "Rule-based triphone mapping for acoustic modeling in automatic speech recognition," in *Text, Speech and Dialogue*. Springer, 2011, pp. 268–275.
- [10] S. Young *et al.*, *The HTK book*. Entropic Cambridge Research Laboratory Cambridge, 1997, vol. 2.

# 3D Editor and Simulator of Petri Nets

Štefan KOREČKO, Oliver NEUBAUER

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

stefan.korecko@tuke.sk

**Abstract** – Petri nets are a popular formal language for specification of non-deterministic and concurrent systems. Their popularity is based on an easy to understand graphical notation. But in the case of more complex systems this advantage is lost as graphs are becoming intricate because of their size and a number of intersecting lines. One of ways to cope with this problem is to move the graph to 3D. In this paper we present a software tool that allows to create, edit and simulate Petri nets in 3D. The tool is based on the OpenSceneGraph 3D graphics toolkit and supports non-traditional input devices, namely a 3D mouse and a data glove.

**Keywords** – Petri nets, 3D editor and simulator, data glove, 3D mouse

## I. INTRODUCTION AND RELATED WORK

Petri nets (PN) [1] are a formal language, introduced by Carl Adam Petri. They evolved from finite automata and have similarly easy to understand graphical notation. But, contrary to automata, they are able to naturally describe concurrent systems with an infinite number of states. They also offer analytical methods, which, for example, allow deriving invariant properties from the structure of the net. PN are typically used for modelling, analysis and simulation of network protocols, operating systems, workflows and other kinds of systems. There are many types of PN with different expressional and modelling power, but all of them have a graphical representation in the form of a bipartite directed graph and behaviour expressed as a token game. The two types of vertices are places (round) and transitions (rectangular). Places hold values, called tokens and these tokens define a state or marking of the net. Marking of the net can be changed by firing of its transitions. A firing of a transition  $t$  can occur when there are enough tokens in its pre-places (i.e. places from which there is an arc to  $t$ ). The firing removes tokens from pre-places of  $t$  and adds tokens to its post-places (i.e. places to which there is an arc from  $t$ ). Number and values of removed and added tokens are defined by expressions written next to arcs.

The easy to understand graphical notation is one of the biggest advantages of PN. But this advantage can be easily lost when a system described by PN is large and complex. Graph of a Petri net is usually represented in 2D and if the net has dozens of vertices it becomes hard to read because of its size and number of intersecting lines. There are several solutions to this problem. One is to use so-called hierarchical PN. In hierarchical PN a net consists of a set of subnets on different levels of abstraction. There is usually one top-level net, which vertices represent sub-systems and connections between them. Each subsystem is then described by a separate subnet. In a subnet, vertices can again represent other subsystems on lower levels or they can be “normal” vertices. A disadvantage of the hierarchical nets is that we never see the system as a whole (on lower levels) and this can make comprehension harder. Another solution that can be used, ideally in combination with the first one, is to add a dimension and allow to create, edit and simulate Petri net models in 3D environment.

There are several existing solutions that combine Petri nets with 3D. The PNVis [2] is an add-on for the Petri Net Kernel [3], which visualizes systems that are modelled as Petri nets in 3D. For 3D models the standard Virtual Reality Modeling Language is used. To be able to perform a visualisation an additional information has to be added to a PN model. This information is about what 3D models are used instead of vertices and arcs and about animations that are played when transitions are fired. Another tool for visualisation of PN models created elsewhere is the BRITNeY suite [4]. It can be used for Coloured Petri nets created in the CPN Tools software [5] and offers various forms of visualisation, including 3D. The last solution we would like to mention is Geist 3D [6], which is a 3D engine based on a Coloured Petri nets-like formalism called Geometry Driven Petri nets (GDPN). GDPN integrate process models, defined as Petri nets, with objects in 3D space, described by means of scene graphs.



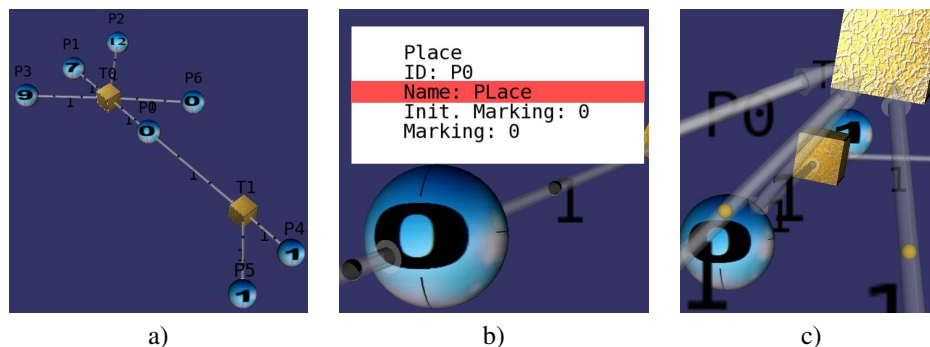


Fig. 1 3D PN editor and simulator: PN graph (a), P0 info editing (b) and tokens during a firing (c).

All the aforementioned solutions offer 3D visualisation of PN simulation but creation and editing of PN happens in 2D. The tool we developed and present in this paper is primarily a 3D editor and simulator of Petri nets. Inspired by tools like PNVis, we also added a possibility to replace the common representation of places and transitions as spheres and boxes by other objects. The tool is based on the OpenSceneGraph 3D graphics toolkit and in addition to the traditional mouse and keyboard also supports a 3D mouse and a data glove, input devices typical for virtual reality applications.

The rest of the paper is organized as follows. The next section briefly describes OpenSceneGraph and explains why this toolkit has been chosen. Section III deals with the editor and simulator itself and section IV discuss the non-traditional input devices support. Finally, section V concludes with plans for future development.

## II. OPENSCEENEGRAPH

OpenSceneGraph (OSG) [7] is an open source high performance 3D graphics toolkit, written in Standard C++ and based on the OpenGL rendering application programming interface. By means of plug-ins it supports most of the commonly used 3D formats. It is multiplatform and can be run under most operating systems available nowadays, including those for mobile platforms, like iOS and Android. This free availability and adaptability was the primary reason for our choice of OSG. Another one is that OSG has a long history of being used in virtual reality (VR) systems, such as VR JuggLua [8] or CalVR [9], so there already is an experience with utilization of the non-traditional input and output devices. OSG is also used in a VR lab at the home institution of the authors.

## III. THE TOOL

The 3D PN editor and simulator prototype, developed by the authors, works with a basic type of Petri nets, called Place transition (PT) nets [1]. The tool is able to load a PT net, written in the Petri Net Markup Language (PNML) [10], display it in 3D, edit and simulate. After loading the net is rendered as in Fig. 1 a): places are spheres, transitions are boxes and arcs are “tubes” with “joints”. The joints are shown as dark grey spheres inside the arcs (best visible in Fig. 1 b)). For every place, transition and arc basic information is shown: names, marking for places and arc expressions for arcs. As the tool supports PT nets the expressions are just natural numbers, called weights, which define how many tokens are consumed or created by firings of adjacent transitions. The tool also allows showing more detailed information about the net parts (in the form of description rectangles above the nodes) and editing it via dialogs as the one in Fig. 1 b). During editing a position of the net parts can be changed in 3D space and we can also choose different colours and textures. In addition, the implicit sphere and box models for nodes can be replaced by other models in the .osg format of OpenSceneGraph. The arcs can be represented as straight tubes, bent tubes or Bezier curves. Most of the settings are available via a 3D menu (Fig. 4). During simulation a fireable transition can be fired by double-clicking it. The actual firing is shown as an animation, which moves tokens along the arcs adjacent to given transition. If arcs are represented as tubes, the movement occurs inside them (small beige spheres in Fig. 1 c)). A user can fire another transition before the animation of a previous firing is completed.

From the implementation point of view the PN model is implemented using the implicit type of 3D objects and scenes representation in OSG, the scene graph. The scene graph is a data structure where parts of an object or a scene are organized in a graph or tree structure. What exactly nodes of a scene graph contain depend on the framework or engine that is implementing the data structure. In our implementation the root node of each object is derived from

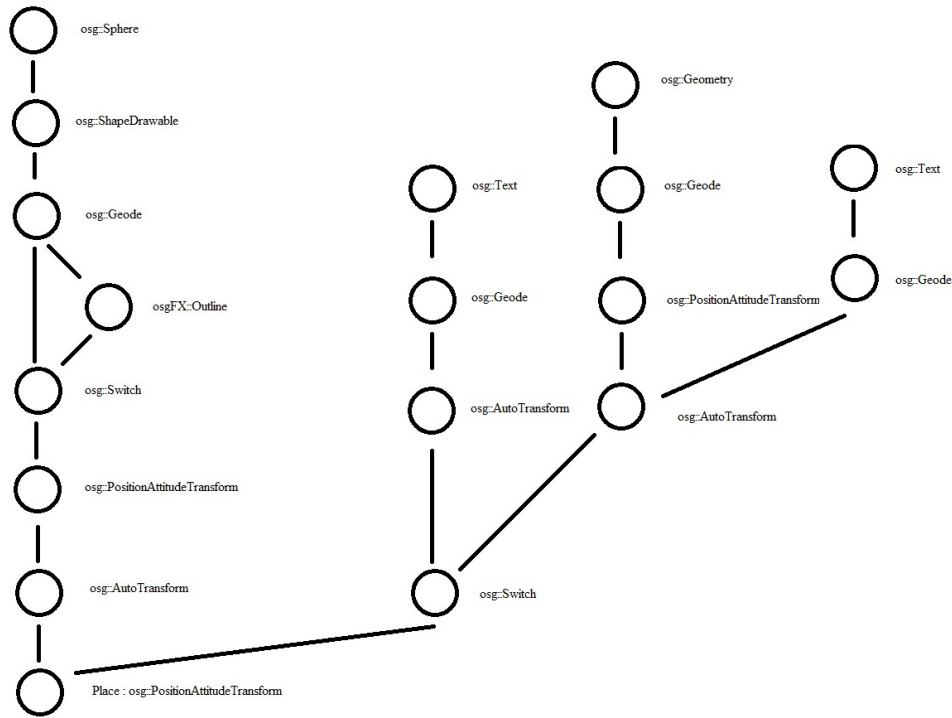


Fig. 2 Hierarchical representation of a place in OpenSceneGraph.



Fig. 3 Supported non-traditional input devices: 3D mouse (a) and data glove (b)

PositionAttitudeTransform class, which provides good support for manipulation with objects and serves as a container for other node classes. For illustration, the scene graph for the PN place is shown in Fig. 2.

The tool also utilizes shaders to apply textures on objects. Transitions and menu objects have two layers of textures. Places contain more layers of textures, depending on the amount of tokens they currently contain. Shaders are also used for creation of quadratic Bezier curves for arcs.

It is not uncommon that a 3D application utilizes some 2D effects. They are used to show objects inside the application facing a designated direction, for example explosions, flares, sky, clouds, and trees. In our tool, if a user rotates the camera around the Petri net model, the description rectangles above PN nodes still face the camera eye. The AutoTransform class was used to implement these 2D effects for description rectangles, 2D text next to the Petri net nodes and for markings in places of the net.

#### IV. NON-TRADITIONAL INPUT DEVICES

The classical input devices such as mouse and keyboard are useable but not ideal for manipulation with objects in 3D space. As one of our ambitions is to integrate the tool into a fully immersive VR environment with stereographic display, we experimented with two of the input devices developed for VR applications.

##### A. 3D Mouse

The first device is the *SpaceNavigator* [11] 3D mouse by 3DConnexion (Fig. 3 a). It allows six degrees of freedom in movement and rotation in 3D space. If we work with a model using

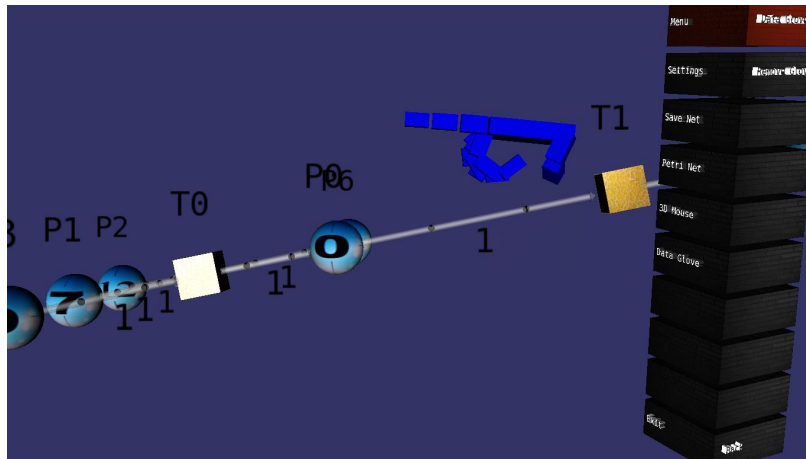


Fig. 4 Virtual hand and menu in the 3D PN editor and simulator

3D mouse, then it feels as if we are holding the model in our hand. The movement and rotation is achieved by (gently) pressing the controller cap of the SpaceNavigator.

There are few options for implementation of SpaceNavigator support into a desired application. One of them is to use VRPN (Virtual-Reality Peripheral Network). VRPN [12] is a set of classes and servers implementing a network-transparent interface between application programs and peripheral devices. Other ones are to use the osgWorks [13] and its library `osgWmX` of matrix operations or the standard development kit (SDK) from 3DConnexion. The last option was chosen for the implementation as the application does not need to support multiple devices and the SDK for a single 3D mouse device is sufficient.

From the user's point of view manipulating a PT net model with SpaceNavigator is smooth and comfortable. The greatest disadvantage of the device is an absence of buttons, so the 3D mouse itself is used just for rotation and movement of the camera or the whole Petri net model. To implement other commands the input from the mouse is combined with the input from the keyboard. An alternative solution is to use a higher 3D mouse model, such as SpacePilot, which has a set of buttons on its device panel.

#### B. Data Glove

The second device is the data glove *DG5-Vhand* (Fig. 3 b) by DG-Tech [14]. The glove uses several sensors for finger bending and rotation. To determine a motion accelerometers are utilized. It is one of the lower-level models, gloves that have more precise sensors and use more advanced technologies (i.e. CyberGlove II) are much more expensive and are used primarily in VR laboratories.

In order to implement the data glove support it is necessary to use the SDK from its producer. Then it is quite easy to data mine exact values of data glove motion. A greater challenge is building up a hand model and creating motion based on the values from the glove. The structure of hand model that represents data glove in application (Fig. 4) was built from basic models of OSG. Scene graph hierarchy proved to be a sufficient solution for the hand model as it provides all the necessary properties for transformations and rotation of the model.

In theory, the data glove should provide more possibilities for control and manipulation with a PN model should be more natural than in the case of the 3D mouse. This anticipation is based on the fact that the glove has separate sensors for finger positions and for hand movement and rotation. So, we can combine gestures, created by fingers with movements of the hand as a whole. However, the practical implementation revealed several issues with reading and interpretation of movement data from the accelerometer of the glove. The problem is that in practice it is hard to rotate a hand but not move it and vice versa. This means that we can have a situation (indicated with a position of fingers), where, for example, a movement is allowed but rotation prohibited. If a user rotates a hand in the situation then the model will not be rotated but the actual rotation will be indicated by the accelerometer and interpreted as a movement.

## V. CONCLUSION

The presented prototype of a 3D editor and simulator of PT nets proved a feasibility of the idea of placing the work with PN models in 3D with a support of VR-oriented input devices such as a 3D mouse and a data glove. Against anticipation, the 3D mouse seems to be more satisfactory input

device as the used data glove model DG5-Vhand. The problem with interpreting accelerometer data correctly in various situations seems to be solvable, however more elaborated algorithms are necessary.

The tool will also benefit from prettier graphical representation and possibility to change the transition firing animation. Implementation of hierarchical PN can be a nice addition, too. The implementation should be performed in such a way that by zooming in we will go seamlessly from top-level nets to more detailed subnets.

#### ACKNOWLEDGMENT

This work has been supported by KEGA grant project No. 050TUKE-4/2012: “Application of Virtual Reality Technologies in Teaching Formal Methods”. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] J. Desel and W. Reisig, “Place/transition petri nets,” in *Lectures on Petri Nets I: Basic Models*, ser. Lecture Notes in Computer Science, W. Reisig and G. Rozenberg, Eds. Springer Berlin Heidelberg, 1998, vol. 1491, pp. 122–173. [Online]. Available: [http://dx.doi.org/10.1007/3-540-65306-6\\_15](http://dx.doi.org/10.1007/3-540-65306-6_15)
- [2] E. Kindler and C. Páles, “3d-visualization of petri net models: Concept and realization,” in *Applications and Theory of Petri Nets 2004*, ser. Lecture Notes in Computer Science, J. Cortadella and W. Reisig, Eds. Springer Berlin Heidelberg, 2004, vol. 3099, pp. 464–473. [Online]. Available: [http://dx.doi.org/10.1007/978-3-540-27793-4\\_27](http://dx.doi.org/10.1007/978-3-540-27793-4_27)
- [3] M. Weber and E. Kindler, “The petri net kernel,” in *Petri Net Technology for Communication-Based Systems*, ser. Lecture Notes in Computer Science, H. Ehrig, W. Reisig, G. Rozenberg, and H. Weber, Eds. Springer Berlin Heidelberg, 2003, vol. 2472, pp. 109–123. [Online]. Available: [http://dx.doi.org/10.1007/978-3-540-40022-6\\_6](http://dx.doi.org/10.1007/978-3-540-40022-6_6)
- [4] M. Westergaard and K. Lassen, “The britney suite animation tool,” in *Petri Nets and Other Models of Concurrency - ICATPN 2006*, ser. Lecture Notes in Computer Science, S. Donatelli and P. Thiagarajan, Eds. Springer Berlin Heidelberg, 2006, vol. 4024, pp. 431–440. [Online]. Available: [http://dx.doi.org/10.1007/11767589\\_26](http://dx.doi.org/10.1007/11767589_26)
- [5] K. Jensen, L. M. Kristensen, and L. Wells, “Coloured petri nets and cpn tools for modelling and validation of concurrent systems,” *Int. J. Softw. Tools Technol. Transf.*, vol. 9, no. 3, pp. 213–254, May 2007. [Online]. Available: <http://dx.doi.org/10.1007/s10009-007-0038-x>
- [6] J. H. Weber-Jahnke and J. Stier, “Virtual prototyping of automated manufacturing systems with geometry-driven petri nets,” *Comput. Aided Des.*, vol. 41, no. 12, pp. 942–951, Dec. 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.cad.2009.06.012>
- [7] (2014) Openscenegraph website. [Online]. Available: <http://www.openscenegraph.org/>
- [8] R. A. Pavlik and J. M. Vance, “Vr jugglua: A framework for vr applications combining lua, openscenegraph, and vr juggler,” in *SEARIS*. IEEE, 2012, pp. 29–35. [Online]. Available: <http://dblp.uni-trier.de/db/conf/searis/searis2012.html#PavlikV12>
- [9] J. P. Schulze, A. Prudhomme, P. Weber, and T. A. DeFanti, “Calvr: an advanced open source virtual reality software framework,” pp. 864 902–864 902–8, 2013. [Online]. Available: <http://dx.doi.org/10.1117/12.2005241>
- [10] (2014) Pnml website. [Online]. Available: <http://www.pnml.org/>
- [11] (2014) Spacnavigator website. [Online]. Available: [www.3dconnexion.com/products/spacnavigator.html](http://www.3dconnexion.com/products/spacnavigator.html)
- [12] (2014) Vrpnl website. [Online]. Available: <http://www.cs.unc.edu/Research/vrpnl/>
- [13] (2014) osgworks website. [Online]. Available: <http://osgworks.vesuite.org/>
- [14] (2014) Dg-tech website. [Online]. Available: <http://www.dg-tech.it/>

# Application of forensic tools to gain description of software behavior

<sup>1</sup>Slavomír Šimoňák, <sup>1</sup>Branislav Madoš, <sup>1</sup>Marek Čajkovský, <sup>1</sup>Ján Hurtuk, <sup>1</sup>Miloš Nagy

<sup>1</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>slavomir.simonak@tuke.sk, <sup>1</sup>branislav.mados@tuke.sk, <sup>1</sup>marek.cajkovsky@tuke.sk,  
<sup>1</sup>jan.hurtuk@tuke.sk, <sup>1</sup>milos.nagy@student.tuke.sk

*Abstract* — This article concerns the creation of solution for introspection of virtual machines with help of forensics tools. The aim of this work is to access to structures located in a virtual machine from the host environment. The work is in the introductory sections dealing with analysis of memory, processes and forensics tools, which can acquire important information from memory dumps. It continues with analysis of virtual machine and method of information acquisition from processes. The next sections shows the proposition and the implementation of the solution. Solution has been tested, results and evaluation of work are written at the end.

*Keywords* — Virtual machine, introspection, memory analysis, forensics tools, memory

## I. INTRODUCTION

The Recency of this topic is undeniable, because the virtualization has become ubiquitous in field of enterprise computing and it is support for cloud computing and data centers. The solution created in this work is focused on virtualization on client side, where the hypervisor is installed on the host operating system as an application. The created solution allows to scan the process memory of virtual machine and find data structures in it. With the help of forensics tools it can read and modify those data structures and by that changing the state of virtual machine. Virtual machine introspection is currently important from a perspective of security. It can be used for monitoring, for example whether the users are running applications approved by company policy. It can also be used for malware detection. The topic of this work is to give some indication of how to access the content of the virtual machine from the hosting environment and to point out the problems in virtual machine introspection.

## II. ANALYSIS

Each process in Windows operating systems has his own virtual address space, which is a set of virtual memory addresses that process can use [1]. For 32 bit operating systems the virtual address space is 4 GB large and for 64 bit its 8 TB [2]. Windows allocates half of this space to processes for their unique private storage and uses the other half for its own protected operating system memory utilization [1]. For 32 bit operating system the lower half of the 4 GB virtual address space is from 0x00000000 through 0x7FFFFFFF and the upper is from 0x80000000 through 0xFFFFFFFF [1]. Virtual machine on client side is represented as an process and so he has its own virtual address space. All processes that are running at operating system inside virtual machine has their own virtual address space. This space is divided into smaller fractions called pages which are then mapped into physical memory assigned to virtual machine, and this physical memory is then mapped into virtual address space of process that represent the virtual machine. There exists two ways to analyze the operational memory of virtual machine. One is to create the memory dump, and analyze it with some forensics tools, and the second is to analyze the user space of a process that represent the virtual machine. Today for memory forensics we are using a lot of sophisticated procedures and tools, which involves localization of different

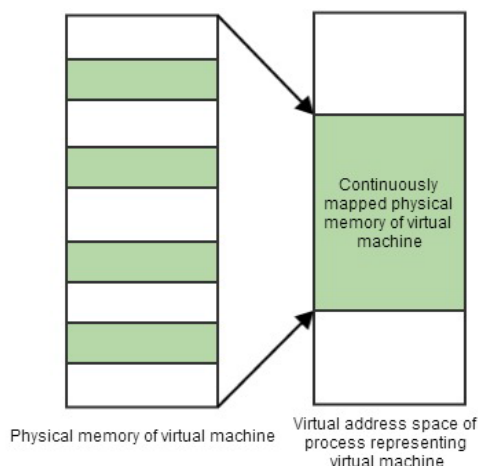
objects in memory as are processes, threads, connections, memory pages [3]. Memory dump is a file that represent the state of physical memory. Memory dumps can be created by many tools, and also hypervisor provide functions that can create \_les with similar content. Forensics tools can be used to analyze the memory dump and provide useful information. However with this approach it is not possible to modify the state of virtual machine memory. Windows operating system has some Windows API functions that can be used to access the content of process memory. Since the whole physical memory assigned to a virtual machine is mapped into the process on host operating system, with this Windows API functions it is possible to get access to content of virtual machine memory. Windows API functions that can be used to access the memory of a process are [1]:

- { ReadProcessMemory() - is used to read memory at specific memory address,
- { WriteProcessMemory() - is used to write data to specific memory address.

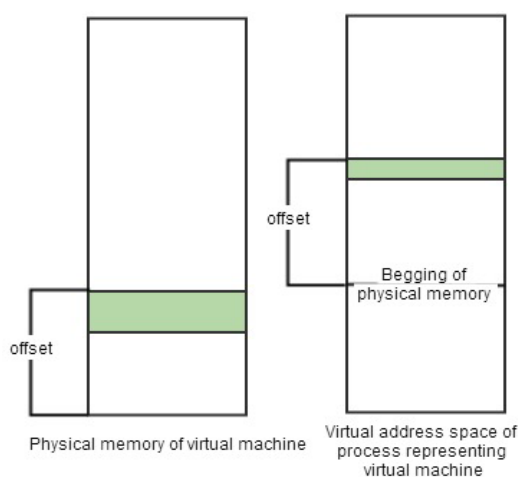
This Windows API functions require virtual address from which to read or write. Hypervisors does not provide information about how is physical memory of virtual machine mapped into virtual memory of process that represent virtual machine. Another problem with live memory analysis is that the content of virtual machine memory is changing with time, and the information acquired from memory may be valid only for short period. When attempting to change the state of virtual machine by using Windows API WriteProcessMemory function it is important to know when and where we can write, because writing to wrong address can cause errors or system failures [4]. It would be ideal if the virtual machine memory wouldn't change during the memory analysis and memory modification. This can be achieved by freezing the state of virtual machine with the pause function that many hypervisors provide.

### III. SOLUTION MECHANISM

This solution is trying to access structures in virtual machine memory, by scanning the memory of a process that represent the virtual machine. Windows API functions are used for reading and writing to process. To access the content of the specific structure in memory, the structure need to be identified in memory. The identification of structure in memory is based on comparing algorithm. First it is necessary to provide some information about searched structure. This information are in a form of expected values of some elements that structure has. For example one element of structure is 4 byte integer and can have value larger than 1000. Then each 4 bytes that holds value greater than 1000 can potentially represent searched structure. The more specific characteristic we provide to the comparing algorithm, the better the chances are to identify the correct structures. However, it is always possible that some data in memory will meet the searching criteria, but they don't actually represent the structure we are looking for, it is just similar to searched structure. Then when the memory is scanned, the read bytes are compared to this expected values. When all the criteria are met, the structure is identified and its memory address is stored. When the address is known we can use Windows API function to read and write to this structure. Another problem comes when the element of structure we are interested in has a reference data type like string. This means that the structure holds address to memory where the object is stored. This address belongs to the operating system inside virtual machine, and unless we have a knowledge how is mapping done between virtual addresses and physical addresses of virtual machine, we cannot use this address to access the object. In order to address this problem, this solution uses forensics tool Volatility that is able to analyze memory dumps and provide information how virtual pages are mapped to physical memory. In this solution we assume that the physical memory of virtual machine is continuously mapped into memory of virtual machine process as it is shown in Figure 1. Then when the volatility provide information at which address is the object located in physical memory of virtual machine, it is necessary to identify the beginning of that physical memory inside process memory from host environment. The procedure to find the start of this memory is described as follows: The memory of the virtual machine process and memory dump are scanned to find one specific object. Then the address of object at process is subtracted by offset of object in memory dump, witch will result in address that represent the start of physical memory of virtual machine as it is shown in Figure 2. The object to find the physical memory start used in this solution is EPROCESS structure of process System for Windows operating system, and task struct structure of process init for Linux operating system. EPROCESS structure represent processes in Windows operating systems, and this structure is different for every version of operating systems. For analyzing Windows operating system this solution support only Windows XP Service Pack 3, for which the search for start is defined.



**Figure 1 Presumed memory mapping of virtual machine**



**Figure 2 Search for start of virtual machine physical memory in virtual machine process.**

Task struct structure does not change between different Linux distribution. However to analyze a memory dump of Linux operating system, the Volatility need explicitly created profile. In this solution I created the profile for operating system Ubuntu 12.04, but it cannot be used for analyzing memory dump of any other Linux distribution, which limits this solution only for analyzing Ubuntu 12.04 Linux distribution. It would be ideal if the object for procedure of identifying the start of the virtual machine's physical memory at virtual machine process would be same between the different version of operating system or between different operating systems. Then with the knowledge where the start of physical memory is and with help of forensic tool Volatility, the solution is able to access even elements of reference data type.

#### IV. TESTING OF THE RESULT

The solution was tested at 32 and 64 bit Windows 8 host operating system. The tested virtual machines used operating system Windows XP Service Pack 3 and Ubuntu 12.04. The assigned memory to virtual machine ranged from 1 to 2 GB. The hypervisors used in testing were VMWare Workstation version 10.0.1 and VirtualBox version 4.3.0. Testing program was created and executed in virtual machine to emulate the searched structure. The testing program created in C# programming language was started inside virtual machine that uses Windows XP operating system. The similar program coded in C++ language was created at virtual machine running Ubuntu. The following text describe part of the test program and its structures:



```
public struct test {
    Int32 mojint1;
    Int32 mojint2;
    strings retazce;
    ...
}

struct strings
{
    string first;
    string second;
}
```

The testing environment was adjusted to eliminate factors that affects the correctness of the results. All the elements of structure were known, the virtual machine was paused and the structure was searched in process of virtual machine. The values stored in first two elements were used for identification of structure. After the structure was identified, the data at the address representing the structure were read and modified. Access to elements of searched structure was tested for both value and reference data types. With access to elements of reference data type it was necessary to have memory dump of virtual machine that will be analyzed by Volatility. Tests proved that structure identification was successful at each tested environment. However some tests identified structures that looked similar to searched structure and modifying their values didn't change output of tested program. As stated earlier in this work it is possible to identify false structures.

## V. CONCLUSION

The goal of this work was to design a solution that will describe how it is possible to acquire information stored in virtual machine memory from hosting environment. Searching and identifying structures in memory is based on comparing algorithm, and the identification is independent of operating system running at virtual machine. Structure identification was successful in all the tests. The forensic tool Volatility was used at accessing the elements of reference type, which analyzed the virtual machine's memory dump. For analyzing a memory dump of Ubuntu operating system, it was necessary to create a profile for Volatility. Analysis of memory dump provide information about how is mapping done between virtual address to physical addresses. This solution is based on assumption that physical memory of virtual machine is continuously mapped into memory of virtual machine process. The testing proved that this assumption is correct only for 64 bit host operating system and with VMWare virtualization platform. Additional research showed that in another environment the physical memory of virtual machine is divided into smaller parts that are not continuously mapped into virtual machine process. Therefore the application is fully effective only for specific environment. This disadvantages might be improved by the next research, and it would be also great to make the searching algorithm more effective or executed in parallel.

## ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and project KEGA 008TUKE-4/2013: Microlearning environment for education of information security specialists.

## REFERENCES

- [1] Pattern matching of partially ordered events in information managment systems / Liberios Vokorokos, Renáta Vokorokosová - 2008. - 1 elektronický optický disk (CD-ROM). In: INES 2008 : 12th International Conference on Intelligent Engineering Systems : Proceedings : February 25-29, 2008, Miami, Florida. - S. l. : IEEE, 2008 P. 169-172. - ISBN 978-1-4244-2083-4
- [2] Mark Russinovich, David A. Solomon, Alex Ionescu: Windows Internals Part 1. sixth edition. Redmount, Washington: Microsoft Press, 2012. 754 p. ISBN: 978-0-7356-4873-9
- [3] Mark Russinovich, David A. Solomon, Alex Ionescu: Windows Internals Part 2. sixth edition. Redmount, Washington: Microsoft Press, 2012. 674 p. ISBN: 978-0-7356-6587-3 Harlan Carvey: Windows Forensic Analysis DVD Toolkit. second edition. Syngerness Publishing, Inc., 2009. 512 p. ISBN 13: 978-1-59749-422-9



- [4] Harlan Carvey: Windows Forensic Analysis DVD Toolkit. second edition. Syngereess Publishing, Inc., 2009. 512 p. ISBN 13: 978-1-59749-422-9
- [5] Chen, P. M., and Noble, B. D. : When virtual is better than real. In Proceedings of the Eighth Workshop on Hot Topics in Operating Systems (HOTOS 01) (Washington, DC, USA, 2001), IEEE Computer Society, p. 133. ISBN:0-7695-1040-X

# Automatizácia generovania požiadavky na zmenu v procese manažmentu zmien

*Martin SARNOVSKÝ, Andrej ZACHARIÁŠ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

[martin.sarnovsky@tuke.sk](mailto:martin.sarnovsky@tuke.sk), [andrej.zacharias@student.tuke.sk](mailto:andrej.zacharias@student.tuke.sk)

**Abstrakt** — Cieľom práce prezentovanej v tomto článku bolo navrhnuť a implementovať spôsob, ako čiastočne automatizovať procesný tok v rámci aktivít vykonávaných v manažmente zmien. Manažment zmien, ako jeden z kľúčových procesov rámca ITIL, je postavený na konceptoch konfiguračnej databázy CMDB. Cieľom práce bolo využiť štruktúru CMDB a extrahovať z nej informácie potrebné pre zostavenie formálneho návrhu zmeny. Tieto informácie boli doplnené o znalosti extrahované z textového popisu požadovanej zmeny, z čoho sme automaticky vygenerovali dokument požiadavky na zmenu. Prístup bol implementovaný a integrovaný s CMDB nástroja OTRS a otestovaný na vzorke testovacích dát.

**Kľúčové slová** — Manažment zmien, manažment IT služieb, ITIL

## I. MANAŽMENT IT SLUŽIEB

Služba svojím spôsobom predstavuje nástroj k poskytovaniu určitej hodnoty zákazníkovi (t. j. používateľovi služby). Hodnota služby pozostáva z toho, čo odberateľovi služby pomôže získať také výsledky, ktoré potrebuje, ale bez toho, aby bol sám zákazník majiteľom špecifických nákladov a rizík spojených s poskytnutím a dodávkou služby. IT služba potom predstavuje špecifický typ služby, ktorého základom je využívanie informačných technológií [1]. Ide teda o explicitne definovanú a opísanú funkcionalitu, poskytovanú informačnými technológiami, ktorá podporuje alebo priamo umožňuje fungovanie rôznych obchodných činností (obchodná činnosť – proces priamo ovplyvňujúci ziskovosť a fungovanie organizácie). Dôležitým faktom je, že pod pojmom IT služba je nutné vidieť aj služby komunikačné. V tomto prípade môžeme hovoriť o explicitnej prítomnosti komunikačných aspektov tým, že sa hovorí všeobecne o informačných a komunikačných technológiách.

Manažment IT služieb (ITSM – IT Service Management) môže byť definovaný ako množina organizačných zručností, ktoré zabezpečujú dodávku kvalitných IT služieb podporujúcich podnikateľské ciele organizácie použitím nákladovo optimálnych prostriedkov. ITSM sa teda zaoberá riadením IT služieb a to tak, aby tieto služby boli poskytované efektívne a kvalitne. ITSM tento cieľ dosahuje pomoci riadenia interakcie medzi ľuďmi, procesmi a technológiami.

Jedným z najčastejšie používaných rámcov ITSM je ITIL. ITIL (Information Technology Infrastructure Library) predstavuje súbor knižných publikácií, ktorý obsahuje zbierku najlepších skúseností (angl. - best practice) z odboru ITSM. Ide o súbor overených konceptov a postupov z praxe, ktoré umožňujú lepšie plánovať, využívať a skvalitňovať využitie informačných technológií zo strany dodávateľov IT služieb, ako aj z pohľadu zákazníka [2]. Aktuálna verzia ITIL (Information Technology Infrastructure Library) V3 zosobňuje najlepšie skúsenosti a osvedčené postupy z praxe, ako riadiť IT služby. ITIL je udržiavaný organizáciou IT Service Management Foundation (itSMF) a predstavuje medzinárodný štandard pre oblasť ITSM. ITIL predstavuje iba návody, ako pristupovať k danej problematike riadenia IT služieb, nepredstavuje normu a čitateľovi poskytuje základný postup na riešenie špecifických situácií, s ktorými sa môže stretnúť dennodenne v praxi. Posledná verzia ITIL V3 bola vydaná v roku 2011 a jej gro je opis celého životného cyklu IT služby [2].

## II. MANAŽMENT ZMIEN PODĽA ITIL

Pred popisom samotného procesu Manažmentu zmien, tak ako ho definuje ITIL musíme definovať niekoľko kľúčových pojmov z tejto oblasti. Konfiguračná databáza (CMDB - Configuration Management Database) je model, ktorý slúži na modelovanie IT infraštruktúry poskytovateľa služieb. Modeluje komponenty infraštruktúry ako konfiguračné záznamy po celú dobu ich životného cyklu. CMDB predstavuje konfiguráciu dôležitých zložiek IT prostredia, ktorej cieľom je pomôcť organizácii poskytujúcej IT služby pochopiť vzťahy medzi jednotlivými zložkami a monitorovať ich konfiguráciu. CMDB je základným komponentom štruktúry ITIL v rámci konfiguračného manažmentu. Prvkami databázy konfiguračného manažmentu v tomto kontexte sú konfiguračné položky (CI, Configuration Items). Konfiguračná položka môže predstavovať ľubovoľný IT komponent ako napr. softvér, hardvér, dokumentáciu, ľudské zdroje, atď. a ich akúkoľvek kombináciu.

Zmena (angl. - change) zahŕňa prídanie, úpravu alebo odstránenie niečoho, čo môže ovplyvniť dodávku IT služby. Ide teda o činnosť, ktorá si vyžaduje zásah do niektorej z konfiguračných položiek v CMDB [3]. Zmeny v IT infraštruktúre môžu vzniknúť ako reakcia na problémy, externe uložené požiadavky, napr. legislatívne zmeny, alebo aktívne v závislosti od snahy dosiahnuť lepšiu efektívnosť a účinnosť poskytovaných služieb. Pre riadenie takýchto zmien obsahuje ITIL (ale aj iné rámce) proces Manažmentu zmien. Úlohou manažmentu zmien je, čo najefektívnejšie riadiť a koordinovať životný cyklus všetkých zmien a zároveň zabezpečiť, aby sa štandardizované metódy a postupy používali pre efektívne a rýchle riešenie všetkých zmien, týkajúcich sa IT infraštruktúr, s cieľom minimalizovať negatívne dopady a narušenie danej služby. Manažment zmien umožňuje zabezpečiť štandardizované metódy, procesy a postupy, ktoré sa používajú pri riadení zmien, uľahčuje účinné a rýchle spracovanie všetkých zmien a napomáha udržiavať správnu rovnováhu medzi potrebou zmeny a potenciálnymi negatívnymi dopadmi zmien. V rámci ITIL-u je manažment zmien súčasťou fázy prechod služieb (Service Transition). Začať proces zmien majú možnosť iba IT tímy prostredníctvom žiadosti o zmenu (RfC – Request for change). Požiadavka na zmenu obsahuje detaily navrhovanej zmeny a môže byť zaznamenaná v papierovej alebo elektronickej forme. Výraz požiadavka na zmenu sa často nesprávne používa na vyjadrenia záznamu o zmene, alebo na vyjadrenie samotného výrazu zmena. Až po zaznamenaní požiadavky na zmenu sa táto vyhodnocuje a evaluuje a v prípade schválenia naplánuje a implementuje. Všetky ostatné požiadavky na zmeny pochádzajú od zákazníkov, užívateľov v organizácii a mimo nej, dodávateľia, atď. vo forme správ (napr. emailom) a v prípade potreby sú tieto požiadavky transformované na RfC [4]. RfC dokument obsahuje informácie ako typ zmeny, dôvod zmeny, implementačné kroky, typ testovania zmien, eskalačný plán, predpokladaná doba trvania, potenciálny vplyv, typ konfiguračných položiek, ktorých sa zmena bude dotýkať, služieb, ktoré budú zmenou postihnuté, atď.. Zodpovednosťou iniciátora zmeny je zabezpečiť dostatočne podrobný popis v žiadosti o zmenu pre manažéra zmeny, aby mohol vykonať jeho kvalifikované posúdenie [7].

### III. NÁVRH GENERÁTORA ŽIADOSTÍ O ZMENU

ITSM podporné softvéry vo všeobecnosti podporujú vytváranie a prácu s RfC dokumentmi. Väčšinou sú implementované základné prostriedky pre spracovanie a trobu RfC dokumentov, ktoré zvyčajne podporujú štruktúru podľa rámca ITIL. Hlavnou ideou práce predstavenej v tomto článku bola automatizácia procesu generovania RfC dokumentov s využitím znalostí, ktoré sú uložené v CMDB a s využitím informácií, ktoré je možné extrahovať zo správy, ktorá predchádza požiadavke na zmenu. Nami predkladané riešenie sme sa rozhodli otestovať na CMDB a ITSM softvéri OTRS [9], z ktorého sme použili aj dátový model CMDB a technológie, na ktorých je toto riešenie postavené.

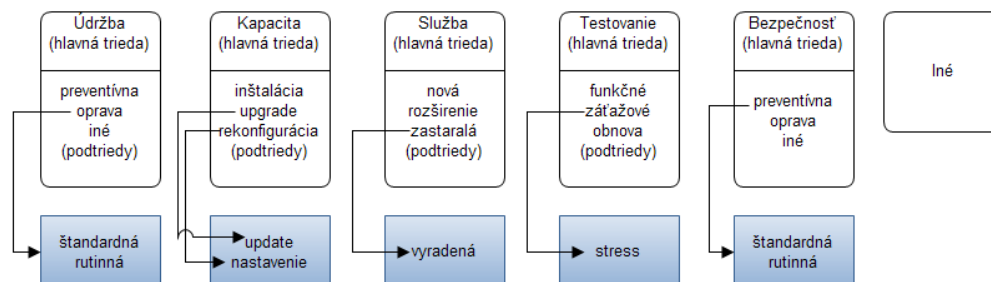
Vzorom šablóny vyrábaného RfC bol oficiálny vzor tohto dokumentu podľa špecifikácie ITIL. Obsahuje všetky potrebné náležitosti potrebné pre formovanie RfC dokumentu. Oficiálnu šablónu sme teda upravili a doplnili o položky, ktoré už nebude za potreby zadávať manuálne. Pre automatizáciu tvorby RfC dokumentu sme využívali 3 druhy údajov:

- Dáta a závislosti odvodené z CMDB
- Dáta extrahované z textového popisu inicializujúceho zmenu

Na základe automatickej extrakcie znalostí z týchto zdrojov je možné automaticky vygenerovať kosru RfC dokumentu s väčšinou informácií, ktoré sú pre formálny návrh zmeny potrebné.

CMDB je úložisko informácií týkajúcich sa všetkých zložiek informačného systému. Dátový model obsahuje hlavnú triedu konfiguračných položiek: Computer a štyri podtriedy: Hardware, Software, Network a Location. Všetky z týchto tried sú popísané viacerými atribútmi. Okrem konfiguračných položiek obsahuje databáza aj informácie o službách, zmluvách ich poskytovania, databázy zákazníkov, používateľov a podpory. CMDB okrem názvov konfiguračných položiek, služieb, atď. obsahuje vzťahy medzi jednotlivými položkami databázy. Relácie sú vo forme „link\_object” a „link\_relation”. Prvá obsahuje názov počiatočného a koncového bodu väzby medzi *Configitem* -> *Service*. Druhá tabuľka obsahuje záznamy o prepojeniach medzi konfiguračnými položkami, resp. službami navzájom. Z hľadiska tvorby RfC dokumentu sú práve tieto väzby kľúčové pre vznik automaticky generovaných častí dokumentu RfC, ktoré tak z ticketu vedú extrahovať všetky dostupné CI, ktoré môže požadovaná zmena postihnúť.

Nasledujúcim krokom je podrobnejšie zistiť z ticketu dôvod plánovanej zmeny. Obr. 1 znázorňuje hlavné kategórie dôvodov zmien s ich podkategóriami tak ako ich definuje ITIL. Medzi hlavné triedy patria Údržba, Kapacita, Služba, Testovanie, Bezpečnosť. Tieto triedy charakterizujú dôvody, kvôli ktorým je potrebné zmenu vykonať z hľadiska zákazníka. Do podtried patria typy dôvodov pre zmeny, ako napr.: prevencia, oprava, inštalácia, rekonfigurácia, rozšírenie, obnova, atď. Implementovaný mechanizmus parsuje správu z ticketu, slová oreže na základný tvar a porovnáva so zostaveným slovníkom, ktorý obsahuje expertne zostavený zoznam slov popisujúci jednotlivé podtriedy. Na základe identifikácie kľúčových slov potom stanoví, ktoré z kategórií k danému popisu zodpovedajú.



Obr. 1 Časová náročnosť vytvorenia modelu pri jednotlivých riešeniach úloh.

#### IV. GENEROVANIE ČASTÍ RFC DOKUMENTU

Táto časť je venovaná ukážke hlavných častí RfC dokumentov pri rôznych dátach získaných zo správy od zákazníka alebo informácií z databázy OTRS. Testy generátora sme vykonali na testovacej CMDB navrhutej pre tieto účely, ktorá popisovala časť fiktívnej IT infraštruktúry a poskytovaných služieb. Testovaciu vzorku hlásení zmien sme použili z testovacej databázy OTRS po preložení do slovenčiny (pre testovacie účely, kvôli databáze kľúčových slov). Na jednotlivých príkladoch si ukážeme generovanie položiek RfC šablóny v konkrétnom prípade. Ticket oznamujúci zmenu v tomto prípade obsahoval explicitne uvedenú požiadavku na update a testovanie aplikačného servera (Windows Server1). Ticket okrem tohoto linku obsahovala správu: “Žiadosť o update a údržbu hardvérových komponentov a operačného systému, o vykonanie záťažových testov routera a siete a o bezpečnostnú opravu v prípade potreby”.

<b>Konfiguracna polozka</b> <input type="checkbox"/> PC_7	<b>Sluzba</b> Sluzba 6	<b>Uroven sluzby</b> SLA 6
<input type="checkbox"/> Monitor 5 <input type="checkbox"/> Router 1 <input type="checkbox"/> Praca <input type="checkbox"/> Siet 2 <input type="checkbox"/> Windows Server 1	<b>Typ sluzby</b> Underpinning Contract	<b>Typ urovne sluzby</b> Transactions

Obr. 2 Časť RfC formuláru s vygenerovanými údajmi z CMDB.

Zo štruktúry CMDB sme odvodili, že aplikačná CI Windows Server1 funguje na hardvéri PC\_7, pripojená Routerom 1 na Siet 2. Na týchto CI funguje časť služby Sluzba 6, ktorá je poskytovaná podľa SLA6. Z modelu infraštruktúry sme teda extrahovali všetky CI, ktoré budú, resp. môžu byť chystanou zmenou ovplyvnené (Obr. 2).

Na základe predošlých údajov a bol text správy rozparsovaný a bola vykonaná analýza kľúčových slov v správe. Na ich základe boli najprv identifikované hlavné znaky zmeny (hlavné triedy) a potom pre každú z nich aj doplnujúce informácie vo forme podtried. Pre uvedený príklad potom vygenerovaná časť formuláru vyzerá takto (Obr. 3).

Nazov zmeny			
<input checked="" type="checkbox"/> Udrzba	<input checked="" type="checkbox"/> preventívna	<input checked="" type="checkbox"/> oprava	<input type="checkbox"/> ine
<input type="checkbox"/> Kapacita	<input type="checkbox"/> instalacia	<input type="checkbox"/> upgrade	<input type="checkbox"/> rekonfiguracia
<input type="checkbox"/> Sluzba	<input checked="" type="checkbox"/> nova	<input type="checkbox"/> rozsirenie	<input type="checkbox"/> zastarala
<input checked="" type="checkbox"/> Testovanie	<input type="checkbox"/> funkčne	<input checked="" type="checkbox"/> zatazove	<input type="checkbox"/> obnova
<input checked="" type="checkbox"/> Bezpecnost	<input checked="" type="checkbox"/> preventívna	<input checked="" type="checkbox"/> oprava	<input type="checkbox"/> ine
<input type="checkbox"/> Ine			

Obr. 3 Časť RfC formuláru s vygenerovanými údajmi z CMDB

## V. ZÁVER

Prezentovaná práca prispieva k zjednodušeniu procesu tvorby RfC dokumentu v procese manažmentu zmien. Hlavnou úlohou bol návrh a implementácia aplikácie ako modulu v systéme OTRS pre automatické generovanie dokumentov RfC z ticketov obsahujúcich žiadosť a popis zmeny. Bol navrhnutý a implementovaný prístup využívania štruktúry CMDB a informácií, ktoré su v nej obsiahnuté a analýza obsahu správy žiadost na základe prítomnosti kľúčových slov. Navrhnutý prístup bol otestovaný na testovacej databáze a na množine príkladov správ so žiadosťami na zmenu.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## REFERENCIE

- [1] R. Schiesser, IT Systems Management. New Jersey, Prentice Hall, ISBN 0-13-087678-X, 2002
- [2] J. van Bon, The guide to IT service management. Addison Wesley, ISBN 0-201-73792-2, 2002.
- [3] ITIL – Service strategy, Office of Government Commerce, ISBN 978-0-11-331045-6, 2007.
- [4] S. Taylor, V. Lloyd, C. Rudd, "ITIL - Service Design", London, TSO, 2007.
- [5] ITIL – Service transition, Office of Government Commerce, 2007, ISBN 978-0-11-331048-7
- [6] ITIL – Service operation, Office of Government Commerce ,2007, [cit.29.03.2011]. ISBN 978-0-11-331046-3.
- [7] D. Nielsen, "How to Control Change Requests", PM Hut, November 15, 2009.
- [8] P. Bereščák, "Vyuzitie ontológií pre oblasť manazmentu IT sluzieb. Košice", diplomová práca, Technická univerzita v Košiciach, 2011
- [9] R. Bakker et. al, OTRS:ITSM – Basic, Documentation, Whitehaven Edition, 2009

# Interaktívna vizualizácia zhukovania textov použitím algoritmu GHSOM v cloudovom prostredí

*Martin SARNOVSKÝ, Zdeno ULBRÍK*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

[martin.sarnovsky@tuke.sk](mailto:martin.sarnovsky@tuke.sk), [zdeno.ulbrik@student.tuke.sk](mailto:zdeno.ulbrik@student.tuke.sk)

**Abstrakt** — Táto práca sa zaoberá problematikou zhukovania textových dokumentov, zhukovacími algoritmami a možnosťou ich vizualizácie. Algoritmus GHSOM (Growing Hierarchical Self Organizing Map) je rozšírením štandardného prístupu k zhukovaniu na báze SOM, ktorý kombinuje adaptivitu rozširovania mapy a hierarchického zhukovania tvorením nových máp. Každá vrstva hierarchickej štruktúry sa skladá z ďalších nezávislých máp, ktoré prispôbujú svoj rozmer vstupným dátam. V tomto príspevku popisujeme návrh vizualizácie modelu GHSOM pre úlohy zhukovania textových dokumentov.

**Kľúčové slová** — Zhukovanie, Text mining, Samoorganizujúce sa mapy

## I. VIZUALIZÁCIA ZHLUKOVACÍCH METÓD

Zhukovanie textov predstavuje proces, kedy sú textové dokumenty zaraďované do tzv. zhukov na základe podobnosti. Každý zhuk obsahuje dokumenty, ktoré sú si čo najviac podobné, a zároveň sú čo najviac odlišné od dokumentov v iných zhukoch. Dokumenty v jednom zhuku si sú obsahovo podobné, preto je možné považovať zhukovanie za prostriedok, ktorý môžeme použiť na zistenie obsahu kolekcie dokumentov. Ide o nekontrolované učenie, učenie sa algoritmov je plne automatické bez zásahu používateľa a počítač hľadá skryté vzory v dátach. Vzhľadom na použitý algoritmus bývajú zhuky najčastejšie disjunktné [1].

Medzi základné modely neurónových sietí, ktoré sú založené na princípe učenia sa bez učiteľa patria samoorganizujúce sa mapy (SOM) [2]. SOM sa často využívajú na zhukovanie textových dokumentov a existuje viacero ich rozšírení. Prvým je algoritmus GSOM (Growing SOM), ktorý umožňuje rozširovanie máp. Toto rozširovanie môže prebiehať dvoma spôsobmi a to pridávaním celých stĺpcov alebo celých riadkov, teda mapa neurónov musí ostať pravouhlého tvaru. Druhým rozšírením je zavedenie viacerých máp, kde každý neurón môže byť expandovaný do novej mapy s rôznym počtom neurónov. Takéto expandovanie môže prebiehať na viacerých úrovniach. Spojením týchto rozšírení vzniká algoritmus GHSOM (Growing Hierarchical Self-Organizing Map) [3]. GHSOM vytvára hierarchickú štruktúru viacerých vrstiev, kde každá vrstva pozostáva z niekoľkých nezávislých SOM.

Pre vizualizáciu jednotlivých modelov a vzniknutých zhukov existuje viacero metód, ktoré sa líšia spôsobom akým vizualizujú nielen samotný proces zhukovania, ale ako podávajú informácie o obsahu jednotlivých zhukov. V tejto časti si uvedieme prehľad tých najpoužívanejších.

Dendrogram je stromový graf, ktorý slúži na obrazovú reprezentáciu hierarchického zhukovania. Podľa typu hierarchického zhukovania objekty spájajú do jedného zhuku alebo rozpájajú jeden zhuk na samostatné objekty. Každý objekt reprezentuje samostatný zhuk, zhuky sa spájajú podľa vzdialenosti medzi sebou **Chyba! Nenašiel sa žiaden zdroj odkazov..** Používa sa pre zobrazenie hierarchických metód zhukovania. Pre vizualizáciu k-stredových metód sa používa zobrazenie na dvojrozmernej mape. Pri viacrozmerných vektoroch vstupných dát musia byť tieto vhodne vhodne pretransformované tak, aby ich bolo možné vizualizovať v dvojrozmernom priestore.

Samoorganizujúce sa mapy nezaraďujú dáta priamo do zhukov, neidentifikujú hranice týchto zhukov ako predošlé metódy zhukovania. Preto je vizualizácia týchto máp kľúčovým faktorom pri analýze dát [5]. Niektoré metódy sa spoliehajú na váhové vektory neurónov, pomocou

ktorých zobrazujú zhluky a hranice medzi nimi (metóda U-Matrix [6]). Ďalší typ metód sa zmeriava na zobrazenie rozloženia vstupných dát, kde zobrazujú na výslednej mape rozloženie všetkých vstupných vektorov alebo len jeden vektor (metódy Hit Histograms a Component Planes) [11].

## II. NÁVRH VIZUALIZÁCIE ALGORITMU ZHLUKOVANIA GHSOM

Vizualizácia modelu GHSOM je možná viacerými spôsobmi v závislosti od spracovaných dát. Vzhľadom na to, že táto práca sa zaoberá zhlukovaním textových dokumentov, navrhli sme model, ktorý je pre túto úlohu vhodný. Navrhnutá metóda vizualizácie spočíva v zobrazení tabuliek, kde každá z nich reprezentuje jeden neurón na mape. Každá tabuľka vizualizuje najvýznamnejšie termy zhuku (zvolené podľa informačného zisku). Jednotlivé mapy sú umiestnené do hierarchického modelu, určeného štruktúrou modelu GHSOM. Po jednotlivých mapách sa dá pohybovať. Práca teda berie inšpiráciu v tabuľkových metódach, no rozširuje ich o zobrazenie informácií o zhlukoch, o štruktúru celej mapy. Používateľovi zobrazuje náhľad na celú štruktúru mapy, na jednotlivé úrovne a dovoľuje mu s mapou interaktívne pracovať.

Využili sme implementáciu algoritmu GHSOM v knižnici JBowl. JBowl (Java Bag of words Library je) open source knižnica určená na dolovanie znalostí z textov. Poskytuje rozhrania na vytváranie aplikácií určených na KDT, obsahuje prostriedky na indexáciu a manipuláciu textových dokumentov, na klasifikáciu a zhlukovanie a možnosti vyhodnotiť niektoré z týchto modelov. Architektúra systému JBowl sa skladá z troch častí [7]:

- Aplikačné programovacie rozhranie – API – sú tu viditeľné triedy, ktoré používateľovi umožňujú prístup k službám na dolovanie v textoch; pre využívanie JBowlu stačí poznať iba toto rozhranie
- Prostriedky dolovania v textoch – ponúka služby dolovania v textoch pre API klientov; môžu byť implementované ako knižnica alebo ako server
- Sklady dolovaných objektov – sklady objektov klasifikácie, zhlukovania, indexovaných dokumentov

Pre metódy vizualizácie bol použitý Processing [9]. Predstavuje open source programovací jazyk a vývojové prostredie určený na vizualizáciu rôzneho charakteru. Jeho úlohou bolo a je zvyšovať gramotnosť programátorov - začiatočníkov, prilákať do sveta programovania umelcov a pomôcť vytvárať vizuálne interpretácie dát, kresliť obrázky, animácie, interakcie jednoducho a rýchlo. Využitie processingu ďaleko presahuje jeho plánované využitie, vzhľadom na možnosť pridávať knižnice, čo umožňuje vytvárať profesionálne vizualizácie z prostredia hudby a dát. Processing je spustiteľný na väčšine operačných systémov Windows, Mac OS a Linux. Vizualizácie vytvorené v tomto jazyku je možné spúšťať priamo v rovnomennom vývojovom prostredí, vo vývojovom prostredí jazyka Java alebo v HTML a JSP stránkach v tagu <canvas>. Pre spúšťanie vizualizácií na webových stránkach je nutná prítomnosť processing.js, čo je Java Script, ktorý prekladá celý skeč do Java Scriptu a je ho možné realizovať na web stránke. Processing má dobre prepracované možnosti získavania dát z externých zdrojov, ako textové dokumenty, xml dokumenty.

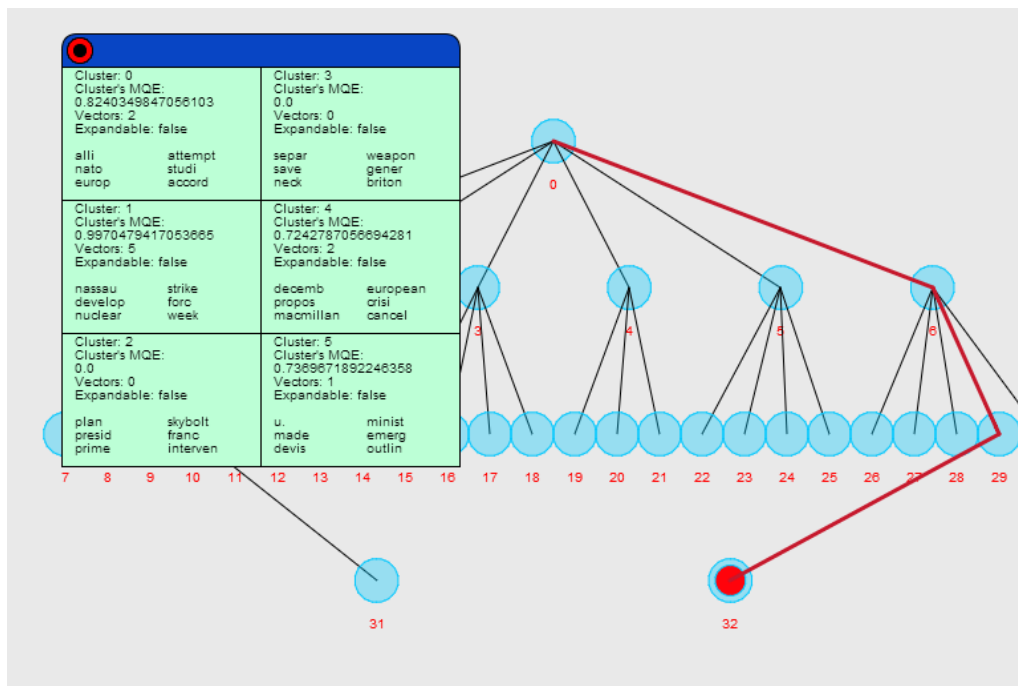
## III. IMPLEMENTÁCIA RIEŠENIA

Metódy knižnice Jbowl vytvárajú mapu prvej úrovne, nájdu na nej expandovateľné neuróny a priradí im potrebné inštancie, z ktorých sa vybuduje celá GHSOM mapa, ktorá sa neskôr napojí na mapu prvej úrovne. Každá čiastková mapa sa vytvára distribuovane použitím GridGain rámca. Po dokončení spájania čiastkových modelov sa vytvorí .xml reprezentácia celej mapy GHSOM. Navrhnutá štruktúra obsahuje všetky potrebné informácie na vytvorenie vizualizácie, Processing vie veľmi ľahko spracovávať .xml dokumenty. Vytvorený a pospájaný model sa načíta zo súboru. Vždy sa načíta jedna mapa a do súboru sa vypíšu informácie o nej. Termy daného zhuku sú zvolené podľa ich informačného zisku v závislosti od termov v ostatných zhlukoch tej istej mapy.

Samotná vizualizácia je implementovaná processing.js (Processing s kombináciou

javascriptu), skladá sa z dvoch hlavných metód, *setup()* a *draw()*. Tieto dve metódy obsahuje každý program vytvorený v jazyku Processing. V prvej metóde sa inicializujú zoznamy máp, neurónov, načítajú sa obrázky, naplní sa zoznam máp a vykreslia sa prvé pozície máp. Druhá metóda beží v nekonečnom cykle a na plátno vykresľuje celú mapu a dva panely.

Interaktívna vizualizácia prebieha po načítaní výsledného modelu celej mapy uloženej v xml. Po načítaní sa vytvorí celková interpretácia modelu GHSOM. Skladá sa z troch častí, z časti vypisujúca informácie o mape, na ktorú sa kliklo, ovládacím panelom a z hlavnej časti, ktorá obsahuje priamo vizualizáciu modelu (Obr. 1).



Obr. 1 Náhľad na hlavnú časť navrhutej vizualizácie.

Interaktivita je reprezentovaná tým, že používateľ má možnosť zasahovať do vizualizácie. Tieto interakcie sú realizované rôznymi akciami, ako napr.:

- pohyb jednotlivých uzlov grafu,
- zobrazenie rôznych druhov informácií (napr. po kliknutí na mapu),
- zobrazovanie / skrývanie počtu úrovní modelu kvôli prehľadnosti,
- zbalenie / rozbalenie uzlov reprezentujúcich mapy,
- zvýraznenie cesty k vybranej mape.

Každý kruh na obrazovke predstavuje jednu mapu, po kliknutí pravým tlačidlom myši na mapu sa zobrazia informácie o mape v ľavom informačnom paneli a v strede obrazovky sa zobrazí detail mapy. Pri zobrazení detailnej mapy sa na pozadí, na celej mape prekreslí cesta k aktívnemu uzlu od hlavnej mapy hrubšou čiarou a mapa sa zvýrazní. Každá mapa sa skladá zo šiestich neurónov (klastrov). Každý neurón má svoje ID, strednú kvadratickú chybu, počet vektorov, informáciu či je expandovateľný (to sa dá zistiť aj podľa zafarbenia pozadia neurónu, červené pozadie znamená expandovateľnosť neurónu) a šesť najdôležitejších termov (Obr. 2). Tieto termy sú zvolené podľa kritéria informačného zisku a ich úlohou je charakterizovať a popísať zhluk, resp. dokumenty, ktoré tento zhluk pokrýva. Na Obr. 2 môžeme vidieť príklad takejto vizualizácie, zobrazená mapa obsahuje neuróny, ktoré ďalej neboli expandované. Každý z nich obsahuje popis počet vektorov dokumentov, ktoré do daného zhluku prislúchajú, termy charakterizujúce dokumenty zhluku, minimálnu kvadratickú chybu. Okrem toho sa v kontextovom okne zobrazuje informácia o mape, resp. neuróne predošlej vrstvy, teda neuróne, ktorého expandovaním vznikla táto mapa. O tomto neuróne máme k dispozícii rovnaké informácie.



Parent Neuron: Parent's map id: 0 Cluster: 4 Cluster's MQE: 1.2418910673833228 Vectors: 17  kill            cong viet            saigon nam            bodi	Cluster: 0 Cluster's MQE: 1.1348353660543556 Vectors: 9 Expandable: false  gaull            konrad adenau          chancellor charl            africa		Cluster: 3 Cluster's MQE: 0.0 Vectors: 0 Expandable: false  eventu            wrong central            threaten congo            poet	
	Cluster: 1 Cluster's MQE: 1.0376550027558669 Vectors: 3 Expandable: false  north            chung unit            kim turkei          korean		Cluster: 4 Cluster's MQE: 0.9832849002453356 Vectors: 3 Expandable: false  himalaya          presenc nehru            delai symbol          lengthen	
	Cluster: 2 Cluster's MQE: 0.0 Vectors: 0 Expandable: false  command        itali bomb            project strateg          mount		Cluster: 5 Cluster's MQE: 0.9362354924817489 Vectors: 2 Expandable: false  soviet            abdel nikita            nasser moscow          execut	

Obr. 2 Vizualizácia konkrétnej mapy spolu s kontextovou informáciou o neuróne vyššej úrovne spolu s doplňujúcimi informáciami.

#### IV. ZÁVER

Cieľom tejto práce bolo navrhnúť a implementovať vizualizáciu algoritmu GHSOM do portálu pre dolovanie dát v textoch, ktorý slúži na výučbové účely. Algoritmus, ktorý je implementovaný distribuovane, je integrovaný do webového portálu pre analýzu textových dokumentov. Navrhnutá interaktívna vizualizáciu tohoto modelu bola integrovaná do existujúceho portálu. Táto vizualizácia poskytuje používateľom možnosť náhľadu na celkovú štruktúru vytváraných modelov, umožňuje vizuálne získavať informácie využiteľné pri riešení úloh dolovania v textoch, vrátane metód vizualizácie rôznych charakteristík modelu, jeho štruktúru, zvolené termy v každej mape. Umožňuje mu sledovať zmenu parametrov algoritmu na výsledný model a spolu s existujúcimi metódami implementovanými v portáli tak poskytuje komplexný nástroj pre riešenie úloh zhukovania textových dokumentov.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### REFERENCIE

- [1] N. O. Andrews, E. A. Fox, "Recent developments in document clustering". Technical report TR-07-35. Department of Computer Science, Virginia Tech, 2007
- [2] T. Kohonen, "Self-organizing maps", Springer-Verlag, Berlín, 1995
- [3] M. Ditttenbach, A. Rauber, D. Merkl, "The Growing Hierarchical Self-Organizing Map", in Proceedings of International Joint Conference on Neural Networks, Como, Italy, 2000
- [4] J. Paralič, et al., "Dolovanie znalostí v textoch", Košice: Equilibria, 2010. 182 s. ISBN 978-80-89284-62-7.
- [5] R. Mayer, T. Aziz, A. Rauber, "Visualising class distribution on self-organising maps." In: Proceedings of the International Conference on Artificial Neural Networks (ICANN'07), volume 4669 of LNCS, Springer. pp. 359–368.
- [6] A. Ultsch, "U\*-Matrix: A tool to visualize clusters in high dimensional data", Department of Computer Science, University of Marburg, Technical Report Nr. 36:1-12.
- [7] P. Bednar, P. Butka, "JBOWL - java bag-of-words library", in 5<sup>th</sup> PhD Student Conference, FEI TU Kosice, Slovakia, pp. 19-20, 2005.
- [8] GridGain Systems: GridGain 3.0 White Paper [online]. 2010. Dostupné na internete: [http://www.gridgain.com/media/gridgain\\_white\\_paper.pdf](http://www.gridgain.com/media/gridgain_white_paper.pdf)
- [9] Processing framework, online: <http://www.processing.org>
- [10] A. Ultsch. Maps for the visualization of high-dimensional data spaces. In Proc. Workshop on Self organizing Maps, Kyushu, Japan, 2003
- [11] J. Vesanto, Data Exploration Process Based on the Self-Organizing Map, PhD thesis, Helsinki University of Technology, 2002

# Optimalizácia distribúcie algoritmu pre tvorbu klasifikačných stromov pre úlohy dolovania v textoch

*Martin SARNOVSKÝ, Jaroslav KOLLÁR*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

[martin.sarnovsky@tuke.sk](mailto:martin.sarnovsky@tuke.sk), [jaroslav.kollar@student.tuke.sk](mailto:jaroslav.kollar@student.tuke.sk)

**Abstrakt** — Cieľom práce bolo navrhnúť optimalizáciu distribúcie algoritmu rozhodovacích stromov pre klasifikáciu textových dokumentov v gridovom prostredí. Optimalizovaný distribuovaný algoritmus budovania rozhodovacích stromov tvorí model paralelne a rozdelenie celkovej úlohy na podúlohy sa vykonáva na základe charakteristík dátovej množiny a vzhľadom na aktuálny stav a vyťaženie používaného gridového prostredia. V tomto článku podrobne popisujeme návrh a implementáciu tohoto riešenia použitím knižnice Jbowl v distribuovanom prostredí GridGain a následne aj vykonané experimenty.

**Kľúčové slová** — Optimalizácia, rozhodovacie stromy, klasifikácia textových dokumentov

## I. DOLOVANIE ZNALOSTÍ V TEXTOVÝCH DOKUMENTOCH

Objavovanie znalostí v množine textových dokumentov (Knowledge discovery in text – KDT, alebo často nazývaný aj Text mining – TM, resp. dolovanie v textoch) sa v dôsledku neštruktúrovanosti jazyka javí ako omnoho zložitejší, než proces objavovania znalostí v databázach [1]. Tento proces môžeme chápať ako špeciálny typ úlohy dolovania znalostí v databázach. Rozdiel môžeme vidieť v tom, že pri databázach pracujeme s dátami v štruktúrovanej forme, v tomto prípade pracujeme s neštruktúrovaným textom. Zvyčajne je možné proces dolovania znalostí v textoch priamo namapovať na proces objavovania znalostí.

V procese klasifikácie textových dokumentov sa pridelujú jednotlivé textové dokumenty do vopred definovaných tried resp. kategórií. Existujú rôzne metódy klasifikácie textov a jednou z nich je aj klasifikácia textov na základe rozhodovacích stromov. Najskôr sa množine tréningových dát sa vytvorí model a potom sa vyhodnotí na testovacích dátach, ktorý sa následne môže používať na reálnych dátach.

Narozdiel od klasických úloh dolovania v textoch je pri klasifikácii textových dokumentov možné vidieť jeden rozdiel. Aj napriek faktu, že sa väčšinou dáta priradujú do jednej kategórie, pri textových dokumentoch sa priradí dokument podľa jeho obsahu nielen do jednej kategórie, ale do skupiny viacerých kategórií [2].

Rozhodovací strom predstavuje pojem, ktorý je veľmi blízky reprezentácii pojmu produkčnému pravidlu, a preto je veľmi ľahké transformovať rozhodovací strom na produkčné pravidlá. Predstavuje reprezentáciu rozhodovacej procedúry pre klasifikáciu príkladov do príslušných tried. Je to grafová štruktúra vo forme stromu obsahujúca koreňový, medziľahlé a listové uzly. Uzly rozhodovacieho stromu reprezentujú triedu alebo testovací atribút. Hrany predstavujú hodnoty testovacieho atribútu. Z každého uzla vychádza toľko hrán, koľko hodnôt má testovací atribút v danom uzle. Na začiatku generovania stromu sa nachádzajú všetky tréningové príklady v koreňovom uzle [3].

Vzhľadom na časovú a výpočtovú náročnosť algoritmov používaných pri riešení úloh na rozsiahlych kolekciami textových dokumentov sa ako možné riešenie tohto problému javí samotná distribúcia a optimalizácia distribúcie týchto algoritmov a dát s využitím SOA – službovo orientovanej architektúry. Technológie pre distribuované objavovanie znalostí sú obzvlášť vhodné pre aplikácie, ktoré zvyčajne spracúvajú rozsiahle objemy dát alebo roztrúsené dátové zdroje, ktoré nie je možné analyzovať v prijateľnom čase na klasických počítačoch.

## II. NÁVRH OPTIMALIZÁCIE DISTRIBUOVANEJ TVORBY KLASIFIKAČNÉHO MODELU

Klasifikácia textov rieši priradenie textových dokumentov na základe ich obsahu do jednej alebo viacerých vopred definovaných tried. Keďže môže byť jeden dokument priradený do viacerých vopred preddefinovaných tried, je potrebné preskúmať aj techniky problému „viacnásobného označovania“. Pri použití stromových klasifikátorov spočíva najvyužívanejší prístup riešenia tohto typu klasifikačného problému v konštrukcii binárnych stromov pre každú kategóriu zvlášť. Tento spôsob riešenia bol navrhnutý a implementovaný v [4] s použitím rôznych iných technológií s využitím gridových služieb. V tomto prípade sa vytvára multi-label rozhodovací strom distribuovane rozdelením čiastkových podúloh na jednotlivé pracovné uzly. Cieľom tejto práce je rozšíriť existujúci algoritmus a zakomponovať do spôsobu distribúcie znalostí o charakteristike spracovávaných dát a aktuálnom stave gridového prostredia.

Navrhnutá logika rozdeľovania úloh spočíva v tom, že sa využijú indexované štatistiky popisujúce dátovú množinu a získajú sa údaje potrebné pre efektívnejšie rozdeľovanie úloh na jednotlivé pracovné uzly. V našom prípade sme použili štatistiku o frekvenciách výskytu jednotlivých kategórií. Tá nám pomôže získať predstavu o jednotlivých čiastkových úlohách (binárnych stromoch pre jednotlivé kategórie), o ich približnej zložitosti a výpočtovej náročnosti. Túto informáciu potom využijeme na rovnomernjšiu distribúciu čiastkových úloh naprieč jednotlivými pracovnými uzlami.

Aktuálny stav monitorovacieho prostredia sme sa rozhodli merať pomocou metód, ktoré poskytuje monitorovací nástroj použitého rámca Gridgain. O jednotlivých pracovných uzloch môžeme zistiť rôzne informácie charakterizujúce ich výkon vrátane informácií o počte procesorov, priemernej záťaži procesora, veľkosti heap pamäť, RAM atď. Tieto informácie, ktoré priamo ovplyvňujú tvorbu modelov na daných uzloch boli zozbierané a bol navrhnutý spôsob ohodnotenia uzlov na základe týchto informácií.

## III. IMPLEMENTÁCIA RIEŠENIA

Použili sme implementáciu stromového algoritmu z Java knižnice JBowl. JBowl (Java Bag of words Library je) open source knižnica určená na dolovanie znalostí z textov. Poskytuje rozhrania na vytváranie aplikácií určených na KDT, obsahuje prostriedky na indexáciu a manipuláciu textových dokumentov, na klasifikáciu a zhukovanie a možnosti vyhodnotiť niektoré z týchto modelov. Architektúra systému JBowl sa skladá z troch častí [5]:

- Aplikačné programovacie rozhranie – API – sú tu viditeľné triedy, ktoré používateľovi umožňujú prístup k službám na dolovanie v textoch; pre využívanie JBowlu stačí poznať iba toto rozhranie
- Prostriedky dolovania v textoch – ponúka služby dolovania v textoch pre API klientov; môžu byť implementované ako knižnica alebo ako server
- Sklady dolovaných objektov – sklady objektov klasifikácie, zhukovania, indexovaných dokumentov

Ako middleware pre clodovú infraštruktúru bol použitý rámec GridGain [6]. Je to middleware (založený na jazyku Java), ktorý slúži na vytváranie distribuovaných aplikácií na cloude, lokálnej sieti. Tieto aplikácie sú väčšinou výpočtovo náročné a pracujú s mohutnými dátami. GridGain sa zameriava na spracovanie dát v reálnom čase, ktoré sú uložené v operačnej pamäti systému, na ktorom je spustený uzol.

Algoritmus distribúcie čiastkových úloh funguje nasledovne:

- Prvým krokom je inicializácia parametrov gridového prostredia (*max heap*, *max RAM*, *max processors*) na nulové hodnoty. Následne sa na hodnotu 0 inicializuje aj premenná uchováajúca celkovú silu uzlov. V cykle sa začne prechádzať uzlami gridu. Z prvého uzla sa načítajú jeho hodnoty parametrov a pretože sú väčšie ako aktuálne maximá, nahradia ich.
- Po ukončení cyklu sa určí absolútne maximum z nameraných hodnôt parametrov. Táto

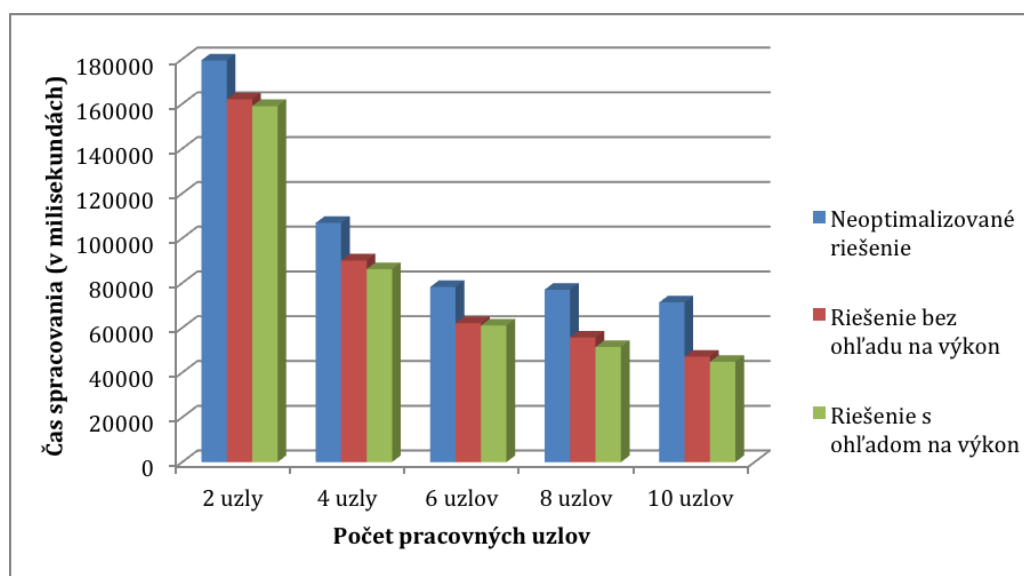
hodnota nám bude slúžiť pre počítanie koeficientov, ktorými budeme uzly ohodnocovať. Koeficient parametra vzniká ako absolútne maximum delené maximálnou hodnotou parametra. Keďže máme pre každý parameter vypočítaný jeho koeficient, môžeme jednotlivé parametre medzi sebou porovnávať a priradiť uzlom ich "silu".

- Prechádzame v cykle všetky uzly a v každom počítame silu uzla ako súčet všetkých parametrov vynásobených koeficientom parametra. Týmto spôsobom je celková sila uzla tvorená tromi hodnotami parametrov vynásobenými ich koeficientmi, ktoré spôsobia vyrovnanie váh jednotlivých parametrov s rôznymi maximami. Pri počítaní síl sa ukladá aj súčet síl uzlov.
- Nasleduje ďalší cyklus ktorý priradí uzlom maximálny počet úloh ktoré majú rátať. To sa deje zistením percentuálnej sily aktuálneho uzla vzhľadom na všetky uzly. Táto hodnota vznikne vydelením sily uzla so súčtom síl všetkých uzlov. Vypočítaný percentuálny podiel vynásobíme počtom všetkých úloh a dostaneme maximálny počet úloh ktoré sa majú na daný uzol priradiť. Tento počet môže byť ale pri finálnom pridelovaní prekročený a to v prípade že sa úlohy nezmestia do žiadneho z uzlov.

Následne sa môžu začať pridelovať úlohy na jednotlivé pracovné uzly zapojené v sieti. Vytvorí sa mapa, do ktorej sa uložia štatistiky o frekvencií výskytu jednotlivých kategórií (vyjadrujúce zložitosti podúloh) a následne sa tieto kategórie zoradia od najfrekventovanejšej po najmenej frekventovanú kategóriu. Na každý uzol sa pridelí určitý počet podúloh podľa toho ako boli uzly ohodnotené a ako sú tieto jednotlivé uzly výkonné. Ak bola priradená uzlu najväčšia sila tak sa na tento uzol pridelí podúlohy výpočtu kategórií s vysokým počtom frekvencií výskytu.

#### IV. EXPERIMENTY

Experimenty boli vykonávané v testovacom prostredí, ktoré tvorilo 12 počítačov HP pripojených na optickú sieť v rámci Technickej univerzity, s konfiguráciou Xeon Processor W3550, 3.07 GHz CPU, 4 GB pamäte RAM, a HDD 450 GB. Hlavným cieľom bolo porovnať neoptimalizované riešenie úlohy tvorby klasifikačného modelu s optimalizovanou verziou. Ďalším cieľom bude nasimulovať zaťaženie pracovných uzlov a sledovať, ako to ovplyvní riešenie úlohy.

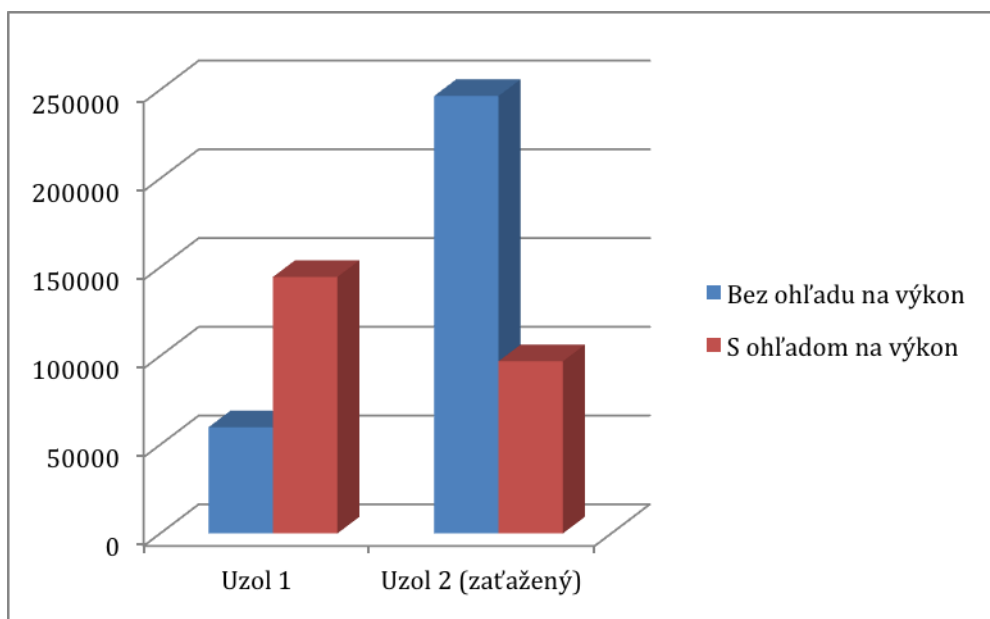


Obr. 1 Časová náročnosť vytvorenia modelu pri jednotlivých riešeniach úloh.

Experimenty prebiehali na dátovej množine Reuters-21578 Modapte, ktorá je zároveň aj vstupom do procesu klasifikácie. V súčasnosti najrozšírejšia kolekcia textových dokumentov používajúca sa v rámci klasifikačných úloh v oblasti výskumu. Je to voľne dostupný korpus dokumentov resp. článkov väčšinou ekonomického charakteru. Dokumenty sú rozdelené na tréningovú a testovaciu množinu. Tréningová množina obsahuje 7769 dokumentov a 28736

termov. Testovacia množina obsahuje 3019 dokumentov. Obidve tie to množiny obsahujú 90 kategórií. Na Obr. 1 môžeme vidieť výsledky vytvárania klasifikačného modelu v závislosti od počtu uzlov v porovnaní s optimalizovanými riešeniami a bez samotnej optimalizácie.

Nasledujúci test sme vykonávali na 2 pracovných uzloch, pričom jeden z nich sme zaťažili. Konkrétne sme zaťažili pracovný Uzol 2 tak, že sme mu vyťažili 2 GB pamäte RAM a procesor. Cieľom bolo nasimulovať zaťaženie jedného z uzlov a otestovať mechanizmus priradzovania čiastkových úloh. Ten by mal na základe navrhnutého modelu odľahčiť zaťažený uzol a priradiť mu jednoduchšie podúlohy na spracovanie. Porovnali sme celkový čas tvorby modelov na zvolených uzloch v prípade, keď sme aplikovali a využili znalosti o stave infraštruktúry so stavom bez tohoto riešenia. Porovnanie je možné vidieť na Obr. 2.



Obr. 2 Časová náročnosť vytvorenia modelu pri jednotlivých riešeniach úloh.

## V. ZÁVER

Cieľom tejto práce bolo optimalizovať distribúciu algoritmu rozhodovacích stromov pre úlohy dolovania znalostí v textoch a tým zvýšiť kvalitu klasifikácie rozsiahlych textových dokumentov. Základ tohto systému predstavuje implementáciu knižnice Jbowl a použitie frameworku Gridgain na vytvorenie distribuovaného prostredia. Takto navrhnutý systém umožňuje paralelný beh celkového procesu, pričom algoritmus na klasifikáciu textových dokumentov môže v tomto prostredí bežať paralelne na rovnakej kolekcii dát, čím sa zvyšuje celková efektivita práce a samotný výkon. Z dôvodu zvýšenia efektivity klasifikácie textových dokumentov sme túto distribúciu algoritmu optimalizovali.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## REFERENCIE

- [1] J. Paralič, et al., "Dolovanie znalostí v textoch", Košice: Equilibria, 2010. 182 s. ISBN 978-80-89284-62-7.
- [2] T. Kačur, "Webová aplikácia pre dolovanie v textoch použitím rozhodovacích stromov", Diplomová práca. Košice: TU FEI, 2013. 71 s.
- [3] K. Machová, "Strojové učenie. Princípy a algoritmy", Košice, Elfa, 2002, 117 s. ISBN 80-89066-51-8.
- [4] I. Janciak, M. Sarnovsky, A. Min Tjoa, P. Brezany, "Distributed classification of textual documents on the Grid". In: High Performance Computing and Communications : Second international conference, HPCC 2006, Munich, Germany, September 13-15, 2006 : Proceedings. Berlin : Springer, 2006. s. 710-718. ISBN 3-540-39368-4.P.
- [5] Bednar, P. Butka, "JBOWL - java bag-of-words library", in 5<sup>th</sup> PhD Student Conference, FEI TU Kosice, Slovakia, pp. 19-20, 2005.

- [6] GridGain Systems: GridGain 3.0 White Paper [online]. 2010. Dostupné na internete: [http://www.gridgain.com/media/gridgain\\_white\\_paper.pdf](http://www.gridgain.com/media/gridgain_white_paper.pdf)
- [7] P. Bednr, P. Butka, J. Paralic, "Java Library for Support of Text Mining and Retrieval" in : Proceedings of Znalosti 2005, s. 162-169, Stará Lesná, 2005.

# Miniaturizovaný systém pre podporu rehabilitácie dolnej končatiny

*Slavomír KARDOŠ, Adam BODZÁS, Peter BALOGH*

Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

slavomir.kardos@tuke.sk

**Abstrakt** — Predmetom príspevku je miniaturizovaný bezdrôtový systém pre použitie v rehabilitačnej praxi. Systém je súčasťou rehabilitačnej pomôcky určenej pre monitorovanie zaťaženia dolnej končatiny v procese liečby. Poskytuje spätnú väzbu a zároveň umožňuje ošetrovujúcemu personálu vyhodnotiť mieru zaťažovania dolnej končatiny pacientom. Systém je zložený z troch nezávislých bezdrôtovo komunikujúcich modulov.

**Kľúčové slová** — ISM, RFM73, Atmel ATmega, rehabilitácia

## I. ÚVOD

Miniaturizované systémy, resp. mikrosystémy (microtransducers) v klasickej i MEMS forme sú v dnešnej dobe neoddeliteľnou súčasťou života človeka. Typicky sú to mikroelektronické obvody, ktoré slúžia na premenu fyzikálnej veličiny na elektrický signál. Obsahujú obvody na spracovanie signálu a výkon funkcií. V ranom štádiu vývoja predstavoval mikrosystém prakticky ľubovoľné elektronické zariadenie s integrovanou elektromechanickou časťou. S rozvojom technológií mikroobrábania sa synonymom mikrosystému stal prvok vytvorený integráciou elektronických obvodov s mikromechanickými prvkami (MEMS). S takýmito prvkami je možné sa stretnúť pri šoférovaní automobilu, pri práci na počítači, pri počúvaní hudby, v medicínskych aplikáciách od sledovania tepu až po komplikované operácie. Technológie výroby takýchto systémov v súčasnosti napredujú súčasne s technológiami výroby integrovaných obvodov, s dôrazom na integráciu, funkcionálnu a spoľahlivosť.

Aplikácia mikrosystémov v medicínskej praxi je podmienená dodržiavaním špeciálnych požiadaviek z hľadiska zapuzdrenia, rozmerov a spotreby energie. Sú to odolnosť voči agresívnym chemickým vplyvom, vrátane krvi a soľných roztokov, miniatúrne rozmery, bezdrôtová komunikácia, veľmi malá spotreba energie, vysoké nároky na teplotnú a driftovú stabilitu a v neposlednom rade implantovateľnosť systému. [1]

## II. BEZDRÔTOVÁ KOMUNIKÁCIA V APLIKÁCIÁCH ELEKTRONICKÝCH SYSTÉMOV

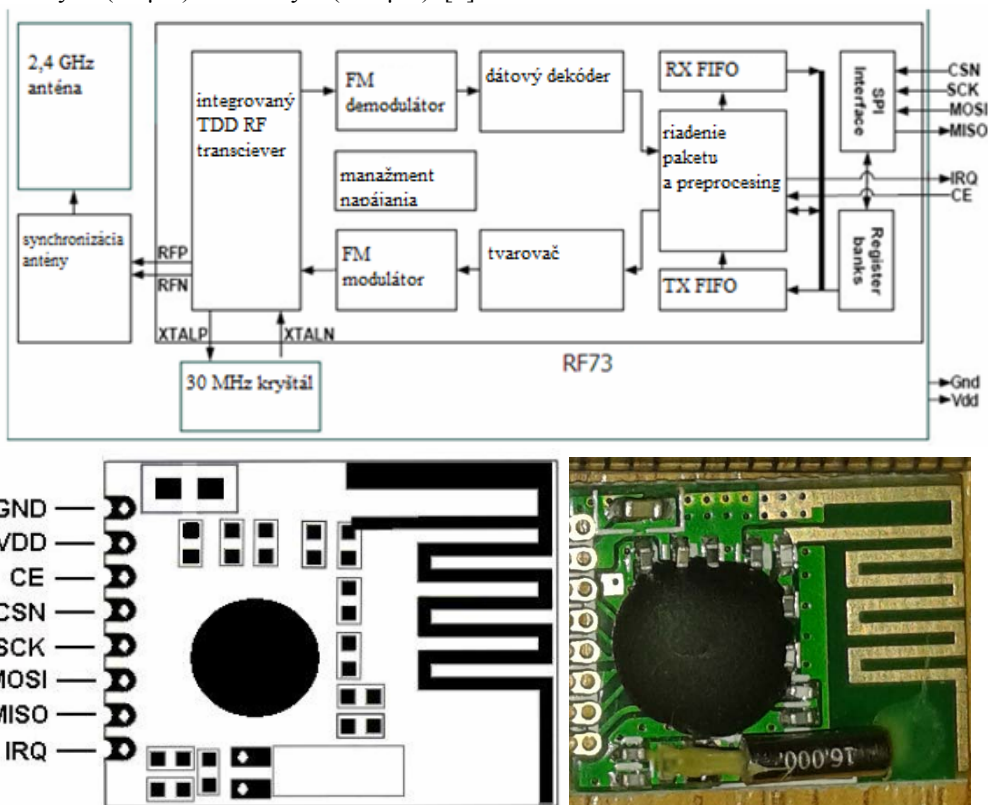
Pre bezdrôtový prenos informácií pre komerčné aplikácie sa v najväčšej miere využívajú technologické štandardy Wi-Fi, Bluetooth, GSM a IrDA. Prenos vo voľných frekvenčných pásmach je stanovený podľa politiky o využívaní frekvenčného spektra platnej v Slovenskej republike. Bezdrôtové (rádiokomunikačné) kanály sú charakterizované šírením rádiových vln. Rádiové vlny sú definované ako elektromagnetické vlnenie frekvencie od 10 kHz až do 3000 GHz. Zodpovedá to vlnovým dĺžkam od 30 km do 0,1 mm. Základné rozdelenie rádiových vln podľa ich vlnovej dĺžky, frekvencie a použitia určuje Rádiokomunikačný poriadok, ktorý je systémom medzinárodných dohôd, ktorý je priebežne aktualizovaný na medzinárodných konferenciách. Na Slovensku je autoritou Telekomunikačný úrad Slovenskej republiky.

Časť frekvenčného spektra je vyhradená pre nelicencované pásmo ISM (industrial, scientific and medical), primárne určené pre vedecké, lekárske a priemyslové aplikácie. Zariadenia majú definovaný maximálny vyžarovaný výkon a neexistuje záruka pre prenos dát bez rušenia. Pásmo je v rozsahu od jednotiek MHz až po GHz a spadajú sem aj technológie Wi-Fi a Bluetooth. [2]

### III. VÝBER RF MODULU V ISM PÁSME

Rádiofrekvenčné moduly sú zariadenia na prenos dát pomocou rádiových vln vo funkcii prijímačov, vysielačov a transceiverov. Pracujú v rôznych ISM pásmach a používajú rôzne typy modulácie. Na trhu existuje nepreberné množstvo rôznych rádiových modulov. Od jednoduchých, ktoré modulujú vstupný signál na nosnú vlnu, až po moduly, ktoré dokážu komunikovať s počítačom, mikrokontrolérom a umožňujú sledovať kvalitu výstupného signálu. Prenos dát je jedným z hlavných problémov, ktoré je potrebné brať do úvahy pri konštruovaní bezdrôtových elektronických systémov. Dôležité je pri prenose dát zabezpečiť ich validitu. Pri prenose na ne pôsobia veľa okolitých vplyvov (interferencia).

Pre realizovanú aplikáciu bol z hľadiska funkcionalít, spotreby, cenovej hladiny a rozmeru použitý bezdrôtový komunikačný transceiver RFM73 pracujúci v pásme 2400 MHz - 2483,5 MHz s moduláciou GFSK. Tento modul dokáže prenášať dáta rýchlosťou 250 kb/s, 1 Mb/s alebo 2 Mb/s. Má programovateľný výstupný výkon, čím sa môže prispôsobiť odoberaný prúd z akumulátora, prípadne z iného zdroja. Umožňuje automaticky generovať obálky signálu na základe prvého naprogramovania. Jeden vysielač podporuje šesť prijímačov. Pracuje od napätia 1,9 V do 3,6 V, jeho riadenie je možné pomocou príslušného mikrokontroléra s komunikačným rozhraním SPI. Rozmery modulu sú 12,8 x 16,8 x 2 mm. Pre efektívne využitie obmedzeného zdroja napájania (batéria) má modul RFM73 tri úsporné módy: Power-down (2,5  $\mu$ A), Standby 1 (50  $\mu$ A) a Standby 2 (330  $\mu$ A). [3]



Obr. 1 Bloková schéma a rozmiestnenie vývodov modulu RFM 73 [3]

### IV. VÝBER RIADIACICH MIKROKONTROLÉROV

Mikrokontrolér (MCU - microcontroller unit) je špeciálny druh mikroprocesora, ktorý obsahuje procesor, pamäť, vstupno-výstupné periférie, režimy spánku vhodné pre batériu napájané aplikácie. Využívané sú hlavne pre jednoduché riadiace aplikácie a sú súčasťou vložených systémov [5]. Mikrokontrolér riadi činnosť celého systému, spracováva a ukladá dáta a spravuje napájací manažment. Z hľadiska programovacích možností, funkcionalít, perspektívy, rozmerov a spotreby boli pre aplikáciu použité 8-bitové mikrokontroléry Atmel ATmega AVR. Pracujú na báze modifikovanej Harvardskej architektúry, obsahujú rozsiahlu výbavu periférií.

Mikrokontrolér ATmega 324pa TQFP je súčasťou konfiguračného modulu. Zvolený bol z hľadiska dostatku vstupno-výstupných vývodov pre klávesnicu, komunikáciu pomocou SPI, externý kryštál, komunikáciu s obvodom FT232RL, externé prerušenia, komunikáciu s modulom RFM 73 a komunikáciu s displejom.

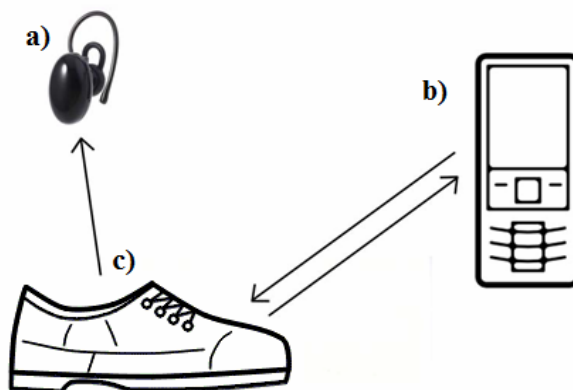
Mikrokontrolér ATmega 328p TQFP je súčasťou monitorovacieho mikromodulu.



Mikrokontrolér ATmega8A je použitý v sluchovom signalizátore pre komunikáciu s modulom RFM 73, ovládacie tlačidlo a signalizáciu preťaženia. [4]-[7]

#### V. KONŠTRUKČNÉ RIEŠENIE SYSTÉMU

Systém je určený na monitorovanie a záznam zaťažovania pohybového aparátu dolných končatín. Jeho súčasťou sú tri samostatné zariadenia. Sluchový signalizátor je určený pre zvukovú signalizáciu nadmerného zaťaženia dolnej končatiny. Monitorovací modul je integrovaný v topánkovej vložke, zaznamenáva zaťaženie a vysiela varovný signál, ak deteguje vopred nastavené preťaženie. Konfiguračný modul je určený na kalibráciu modulu v topánke, následné čítanie zaznamenaných dát a ich prenos. Nabíjanie akumulátorov v bezdrôtových moduloch je riešené použitím autonómneho nabíjacieho obvodu MAX1555EZK+ v puzdre SOT-23-5. Obvod FT232RL slúži ako rozhranie z USB na USART.



Obr. 2 Ilustrácia komunikačných smerov systému - a) sluchový signalizátor, b) konfiguračný modul, c) monitorovací modul integrovaný v topánkovej vložke so snímacím systémom

#### Sluchový signalizátor

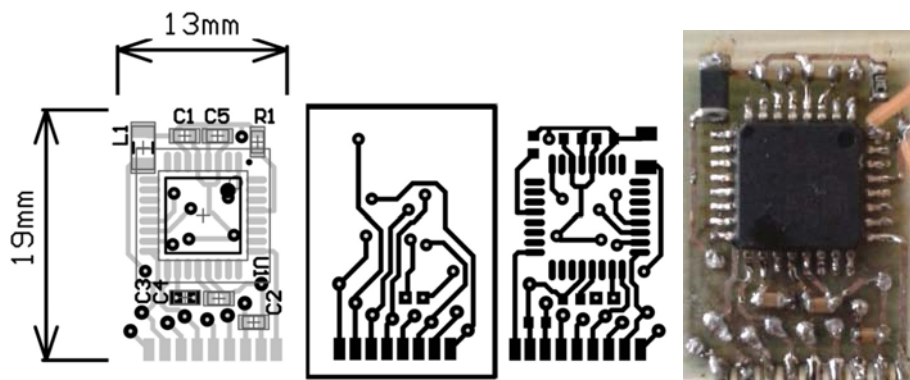
Sluchový signalizátor slúži pre zvukovú signalizáciu nadmerného zaťaženia spodnej končatiny. Obsahuje RF vysielač modulu RFM 73, mikrokontrolér ATmega8, LED diódu, USB konektor micro USB, Li-pol akumulátor, reproduktor, nabíjací obvod MAX1555CZK+ a stabilizátor napätia 3,3 V. Obvod obsahuje tlačidlo pre uvedenie systému do režimu hlbokého spánku a prebudenie. Mikrokontrolér ATmega8 slúži na spracovanie prijatého signálu a vykonávanie hlavných riadiacich funkcií.



Obr. 3 Sluchový signalizátor a) v puzdre z Bluetooth handsfree, b) pred zapuzdrením

#### Monitorovací mikromodul

Monitorovací mikromodul je určený na sledovanie, zber, spracovanie a bezdrôtový prenos informácií o zaťažení končatiny. Senzorickú časť systému tvorí odporový tenzometer. Jeho výstupom je analógové napätie, ktoré je privedené na A/D prevodník mikrokontroléra ATmega328p. Obvod v nekonečnej slučke vzorkuje dáta z A/D prevodníka. Po prekročení určitej hranice mikrokontrolér vyšle signál do sluchového signalizátora, ktorý upozorní pacienta na nadmerné zaťaženie a uloží do pamäte informáciu o zošliapnutí i preťaženi.



Obr. 4 Topológia monitorovacieho mikromodulu a osadený modul

### 1.7 Konfiguračný modul

Funkciou konfiguračného modulu je bezdrôtovo konfigurovať monitorovací mikromodul, čítať dáta a prenášať ich cez USB rozhranie do PC. Pomocou konfiguračného modulu je možné kalibrovať viacero monitorovacích mikromodulov. Každý monitorovací mikromodul môže obsahovať jedinečný ID kód. Využité bolo puzdro, zobrazovacia jednotka a akumulátor mobilného telefónu Nokia 3310. Úvodné logo bolo navrhnuté programe fast LCD. Riadenie displeja s 84 x 48 obrazovými bodmi je realizované pomocou rozhrania SPI. Pre klávesnicu s 15-timi funkčnými klávesmi je použitá koncepcia s kovovými mostíkmi (metal dome). Prepojenia na DPS boli vyhotovené pomocou polymérnej vodivej pasty na báze striebra.



Obr. 5 Konfiguračný modul po prepojení vrstiev a osadení, klávesnicová fólia s kovovými mostíkmi a kompletný modul

## VI. ZÁVER

Počas riešenia konštrukcie systému pre medicínsku aplikáciu bol kladený dôraz na výber vhodnej súčiastkovej základne z hľadiska funkcionalít, integrácie, spôsobu ovládania, typu bezdrôtovej komunikácie, spotreby energie a ďalších. Aplikácia je prototypom, ktorý je potrebné testovať v štandardných podmienkach zaťažovania dolnej končatiny v procese rehabilitácie. Pôsobia na ňu okolité vplyvy elektromagneticky zarušeného prostredia, ako aj klimatické vplyvy. Aplikácia je naprogramovaná ako jeden pár zariadení a univerzálny konfiguračný modul. Systém je možné vylepšiť o generátor reálneho času a softvérové rozhranie pre pokročilé spracovanie dát. (VEGA č. 1/0059/12)

## LITERATÚRA

- [1] SLOSARČÍK, Stanislav – Kardoš, Slavomír – JURČIŠIN, Michal: Výrobné procesy v elektronike 2. Viena, s.r.o., 125 s. - ISBN 978-80-8126-096-4
- [2] ČANDA, Pavel: Bezdrátový přenos dát v pásmu ISM [cit. 4. 4. 2014] Dostupné na internete: <https://dSPACE.vutbr.cz/xmlui/bitstream/handle/11012/7233/MMSE.pdf?sequence=2>
- [3] RFM73 datasheet, [cit. 9.1.2014] Dostupné na internete: [http://www.hoperf.com/upload/rf/rfm73\\_datasheet.pdf](http://www.hoperf.com/upload/rf/rfm73_datasheet.pdf)
- [4] HRABÁČEK, Jiří: Komunikace mikroprocesoru s okolím - 2.díl, 2002, ISBN 80-86056-73-2
- [5] Microcontrollers [cit. 14.3.2014], Dostupné na internete: <http://microcontroller.askdefine.com/>
- [6] Embedded products [cit. 14.3.2014], Dostupné na internete: <http://www.embedded.com/products/all>
- [7] AVR Microcontrollers [cit. 1.1.2014], Dostupné na internete: <http://www.atmel.com/>

# Rozpoznávanie hovoriaceho z emocionálnej reči na princípe i-vektorov

*Lenka MACKOVÁ, Anton ČIŽMÁR*

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky, Technická univerzita v Košiciach, Slovenská republika

lenka.mackova@tuke.sk, anton.cizmar@tuke.sk

*Abstrakt* — V poslednom čase sa prístup použitia i-vektorov v procese verifikácie hovoriaceho stal veľmi úspešný a populárny. Princíp použitia i-vektorov je založený na myšlienke reprezentácie zvukového záznamu málo rozmerným vektorom pevnej dĺžky. V literatúre sa v systéme rozpoznávania hovoriaceho na báze i-vektorov väčšinou využívala telefónna reč doplnená nahrávkami z iného zdroja. V tomto experimente boli použité nahrávky emocionálnej reči mužských subjektov získaných z našej vlastnej emočnej databázy SUS v slovenskom jazyku. V procese tréningu modelov hovoriacich boli použité dva rôzne princípy. V procese extrakcie príznakov bol extrahovaný rôzny počet mel frekvenčných keprálnych koeficientov (Mel Frequency Cepstral Coefficients) s logaritmickou energiou, delta, akceleračnými a terciárnymi regresnými koeficientmi. V procese klasifikácie sa použila Mahalanobisova skórovacia metóda s radiálnou normalizáciou na základe vlastného faktora (Eigen Factor Radial Normalization) a metóda skórovania na základe kosínusovej podobnosti (Cosine Similarity Scoring) s medzitriednou kovariančnou normalizáciou (Within-class Covariance Normalization).

*Kľúčové slová* — verifikácia hovoriaceho, i-vektory, totálna variabilita, emócie.

## I. ÚVOD

V existujúcich štúdiách boli testované rôzne prístupy systémov verifikácie hovoriaceho [1]. Najpopulárnejšie systémy vychádzajú z generatívnych modelov, ako sú napríklad Gaussove modely zmesí (Gaussian Mixture Models - GMM) založené na tréningu univerzálneho modelu pozadia (Universal Background Models - UBM) [2], metóda vlastných hlasov (eigenvoices) [3] a vlastných relácií (eigenchannels) [4]. Do kategórie GMM-UBM modelov patrí aj doteraz najefektívnejší nástroj verifikácie hovoriaceho – analýza spojených faktorov (Joint Factor Analysis - JFA) [5].

JFA pracuje s modelovaním variability hovoriaceho a kompenzácie variability relácie, pričom využíva princípy veľkorozmerných GMM supervektorov. Experiment, ktorý bol vykonaný v [10] ukázal, že v prípade použitia JFA sa v relačnom priestore nachádzali aj informácie hovoriaceho. Na základe tohto experimentu Dehak [8] navrhol extraktor príznakov, ktorý pracuje s málo rozmerným podpriestorom GMM supervektorov. Tento priestor sa nazýva priestor totálnej variability a reprezentuje ako variabilitu hovoriaceho, tak variabilitu relácie.

Vzhľadom na to, že priestor totálnej variability obsahuje relačné informácie, systém pracujúci na báze i-vektorov využíva rôzne kompenzačné techniky ako sú lineárna diskriminantná analýza (Linear Discriminant Analysis - LDA) [7], medzitriedna kovariančná normalizácia (Within-class Covariance Normalization - WCCN) a projekcia rušivých atribútov (Nuisance Attribute Projection - NAP).

Tento experiment je zameraný na zistenie vplyvu použitia emočnej reči v procese rozpoznávania hovoriaceho.

Na extrakciu i-vektorov bol použitý voľne dostupný programový balík určený na rozpoznávanie hovoriaceho – Alize/LiaRal [11]. V tomto balíku sú implementované viaceré klasifikačné metódy – skórovanie na základe kosínusovej podobnosti (Cosine Similarity Scoring - CSS), Mahalanobisova skórovacia metóda, dvojkovariančné skórovanie a PLDA skórovanie. V experimente boli použité dve klasifikačné metódy, Mahalanobisova klasifikácia a CSS skórovacia metóda.

Tento článok je organizovaný nasledovne. V kapitole 2 sú opísané princípy extrakcie i-vektorov, kapitola 3 je venovaná klasifikačným technikám použitým v tejto práci. Použité

normalizačné techniky sú opísané v kapitole 4. Kapitola 5 predstavuje SUS emocionálnu databázu. Kapitola 6 sa venuje samotnému experimentu a kapitola 7 je záverečnou kapitolou.

## II. PROCES EXTRAKCIE I-VEKTOROV

JFA [6] je založená na dekompozícii GMM supervektora, ktorý je závislý na hovoriacom, na separátne informácie týkajúce sa hovoriaceho a relácie. JFA teda reprezentuje záznam hovoriaceho ako supervector  $M$ , ktorý je definovaný ako

$$M = m + Vy + Ux + Dz, \quad (1)$$

kde  $m$  je supervektor nezávislý na hovoriacom alebo prenosovom kanále,  $D$  je diagonálna reziduálna matica,  $V$  je eigenvoice reziduálna matica. Obe matice  $V$  a  $D$  reprezentujú priestor informácie o hovoriacom. Matica  $U$  je eigenchannel matica, ktorá definuje priestor relácie. Vektory  $y$  a  $z$  sú faktory závislé na hovoriacom. Vektor  $x$  je faktor závislý na relácii. Každý z týchto vektorov je náhodnou premennou s normálnou distribúciou  $N(0; I)$ .

Hlavná myšlienka použitia i-vektorov spočíva v definovaní jediného priestoru nazývaného totálna variabilita. Tento priestor zahŕňa informácie o hovoriacom a prenosovom kanáli súčasne. GMM supervector nezávislý na hovoriacom a relácii môže byť tak prepísaný z rovnice 1 do tvaru

$$M = m + T\omega, \quad (2)$$

kde  $m$  je supervektor závislý na hovoriacom a relácii,  $T$  je rektangulárna matica nízkeho rádu a  $\omega$  sú faktory nazývané i-vektory.

Supervektor  $M$  má normálnu distribúciu s priemerným vektorom  $m$  a kovariančnou maticou  $TT^t$ . Proces tréningu matice totálnej variability sa dá porovnať s tréningom matice vlastných hlasov (eigenvoice). Rozdiel je v tom, že pri tréningu eigenvoice matice sa nahrávky hovoriaceho považujú za nahrávky patriace tej istej osobe. V procese tréningu matice totálnej variability sa celý súbor nahrávok hovoriaceho považuje za nahrávky patriace rôznym subjektom.

## III. METÓDY SKÓROVANIA V SYSTÉME VERIFIKÁCIE HOVORIACEHO

### A. Mahalanobisova klasifikačná metóda

Mahalanobisova klasifikácia je jednou z metód použitých v tomto experimente. Je založená na porovnávaní tréningového súboru i-vektorov s triedou i-vektorov už známych hovoriacich. Cieľom tejto metódy je nájsť triedu, do ktorej bude sledovaný i-vektor patriť. Aplikáciou Bayesovho riešenia sa získa najpravdepodobnejšia trieda, do ktorej bude zaradený sledovaný i-vektor. Priradenie i-vektora  $w$  hovoriacemu  $s$  je dané rovnicou

$$(w - \bar{w}_s)^t W^{-1} (w - \bar{w}_s) = \|w - \bar{w}_s\|_{W^{-1}}^2, \quad (3)$$

kde  $\bar{w}_s$  je priemer (centroid) triedy  $s$  a  $W$  je medzitriedna kovariančná matica definovaná rovnicou (7). Vychádzajúc z Mahalanobisovej metriky matice  $W^{-1}$  element  $w$  bude priradený priemeru najbližšej triedy.

Skóre dvoch i-vektorov  $w_1$  a  $w_2$  je dané logaritmicou pravdepodobnosťou toho, že skórované i-vektory patria do triedy danej kovariančnou maticou  $W$ . Centroidom takejto triedy bude jeden z i-vektorov alebo ich priemer. Výsledná Mahalanobisova skórovacia rovnica je zapísaná v tvare

$$\text{score}(w_1, w_2, \cdot) = -\|w - \bar{w}_s\|_{W^{-1}}^2. \quad (4)$$

### B. Skórovanie na základe kosínusovej podobnosti

Princípom klasifikačnej metódy CSS je skórovanie dvoch i-vektorov na základe porovnávania kosínusovej vzdialenosti týchto i-vektorov. V prípade CSS sa výsledné i-vektory vypočítavajú z ich vstupných dát vo vstupnej fáze, preto v tomto prípade nie je potrebný žiaden cieľový model. V testovacej fáze, sa skóre dvoch i-vektorov počíta na základe kosínusovej vzdialenosti medzi testovacím a cieľovým vektorom. Kosínusové skóre je definované ako

$$\text{score}(\hat{w}_1, \hat{w}_2, \cdot) = \frac{\hat{w}_1 - \bar{w}_{imp} \quad \hat{w}_2 - \bar{w}_{imp}}{\|C_{imp} \hat{w}_1\| \|C_{imp} \hat{w}_2\|}, \quad (5)$$

kde  $\bar{w}_{imp}$  je priemer falošných subjektov a  $C_{imp}$  je druhá odmocnina kovariančnej matice týchto subjektov. Výsledné skóre je následne porovnávané s rozhodovacím prahom  $\theta$ .

Nevýhodou tejto klasifikačnej metódy je to, že CSS uvažuje pri porovnávaní dvoch i-vektorov iba ich uhol a nie ich absolútne veľkosti. Pre zlepšenie skórovania a odstránenie rozdielu magnítud porovnávaných i-vektorov sa používa normalizačná technika WCCN.

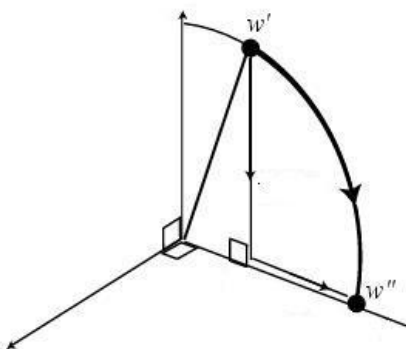
#### IV. NORMALIZAČNÉ TECHNIKY

Vzhľadom na to, že extrahované i-vektory obsahujú informácie o hovoriacom aj informácie relácie, na kompenzáciu relačných informácií sa používajú normalizačné techniky. V tomto experimente boli použité dve normalizačné techniky, WCCN a EFR. Tieto techniky pripravujú extrahované dáta pre ich ďalšie spracovanie klasifikačnými metódami.

##### A. Radiálna normalizácia na základe vlastného faktora

Radiálna normalizácia na základe vlastného faktora sa zvykne nazývať tiež radiálnou NAP [12]. Základom NAP techniky [13] je odhad variability relácie z medzitriednej kovariančnej matice.

Princípom EFR je potlačenie neprípustnej dimenzie a následná normalizácia vektorov. V ortogonálnom priestore to znamená rotáciu každého vektora  $w$  k prvej základnej osi v priestore relácie ako je znázornené na Obr. 1.



Obr. 1 Princíp radiálnej NAP

V metóde EFR je priemer vektora označený ako  $\bar{w}$  a  $V$  je kovariančná matica trénovacieho súboru i-vektorov. Normalizácia i-vektorov a redukcia variability relácie je definovaná rovnicou

$$w = \frac{\frac{1}{\sqrt{2}}(w - \bar{w})}{\sqrt{(w - \bar{w})V^{-1}(w - \bar{w})}}. \quad (6)$$

Medzitriedna kovariančná matica sa vypočíta nasledovne

$$W = \sum_{s=1}^S \frac{n_s}{n} W_s = \frac{1}{n} \sum_{s=1}^S \sum_{i=1}^{n_s} (w_i^s - \bar{w}_s)(w_i^s - \bar{w}_s)^t. \quad (7)$$

V tomto prípade je  $W_s$  kovariančnou maticou hovoriaceho  $s$ ,  $n$  je počet kompletného súboru nahrávok,  $n_s$  je počet nahrávok patriacich subjektu  $s$ ,  $w_i^s$  sú trénovacie i-vektory hovoriaceho  $s$  a  $\bar{w}_s$  je priemer týchto i-vektorov.

##### B. Medzitriedna kovariančná normalizácia

WCCN prístup je aplikovaný v systémoch podporných vektorov (Support Vector Machines – SVM) systémoch. Hlavnou myšlienkou tohto princípu je separácia dát na cieľové dáta hovoriaceho subjektu a dáta falošných subjektov. Testovacie príznaky sa potom porovnávajú so všetkými ostatnými príznakmi. Z dôvodu minimalizácie chyby falošných akceptácií a falošných zamietnutí v procese trénovania sa definuje súbor prahov klasifikačnej metriky. Výsledok je následne produkovaný lineárnym jadrom vo forme rovnice

$$k(w_1, w_2) = w_1^t R w_2, \quad (8)$$

kde  $R$  je symetrická, pozitívna, polodefinovaná matica. Optimálna kernelová matica je definovaná ako  $R=W^{-1}$ , kde  $W$  je medzitriedna kovariančná matica vypočítaná z príznakov všetkých falošných subjektov. Pre potreby rozpoznávania hovoriaceho je medzitriedna kovariančná matica  $W$  vypočítaná nasledovne

$$W = \frac{1}{S} \sum_{s=1}^S \frac{1}{n_s} \sum_{i=1}^{n_s} (w_i^s - \overline{w_s})(w_i^s - \overline{w_s})^t. \quad (9)$$

V tomto prípade je  $S$  výsledný počet hovoriacich,  $n_s$  je počet nahrávok subjektu  $s$  a priemer  $i$ -vektorov každého hovoriaceho je daný ako

$$\overline{w_s} = \frac{1}{n_s} \sum_{i=1}^{n_s} w_i^s. \quad (10)$$

Matica WCCN označená ako  $B$  sa vypočíta prostredníctvom Choleskyho dekompozície matice  $W^{-1}=BB'$ .

## V. EMOČNÁ DATABÁZA

Pre účely tohto experimentu bola vytvorená nová emočná databáza SUS. Na vytvorenie tejto databázy sa použili nahrávky voľne dostupných televíznych relácií v slovenskom jazyku. Databáza pozostáva z približne 1900 mužských a ženských nahrávok v emóciách hnevu, agresivity, zvedavosti, prekvapenia, znechutenia, sklamania a neutrálneho stavu.

Všetky nahrávky databázy boli vzorkované vzorkovacou frekvenciou 16 kHz a dĺžka každej nahrávky je približne 6 sekúnd. Celá databáza má trvanie približne tri hodiny.

V tomto experimente boli použité emočné nahrávky troch mužských hovoriacich subjektov v emóciách neutrálneho stavu a zvedavosti. Dôvodom pre použitie iba dvoch emócií je dostatočný počet nahrávok v týchto emóciách pre zodpovedajúcich hovoriacich, z ktorých bol vytvorený plnohodnotný tréningový a testovací súbor s dostatočným počtom nahrávok.

## VI. EXPERIMENT A VÝSLEDKY EXPERIMENTU

Tento experiment bol rozdelený do dvoch častí. V prvej časti sa na tréning modelu jednotlivých subjektov použili všetky nahrávky z emočnej databázy SUS v emóciách neutrálneho stavu a zvedavosti. V druhom prípade sa na tréning modelu hovoriaceho použila iba jedna emócia. Hlavnou myšlienkou tohto postupu bolo overenie výsledku rozpoznávania hovoriaceho pri tréningu zmiešaného a nezmiešaného modelu hovoriaceho v prípade aplikácie  $i$ -vektorov.

V prvej časti experimentu bolo na tréning univerzálneho modelu pozadia (UBM) použitých 247 nahrávok databázy SUS. V tréningovej fáze bolo použitých 297 a v testovacej fáze 175 nahrávok. Výsledkom tréningovej fázy v tomto prípade boli tri modely, ktoré vznikli natréningovaním z nahrávok oboch spomenutých emócií.

V druhom prípade ostal počet nahrávok na tréning modelu pozadia a testovacích nahrávok rovnaký. Počet tréningových nahrávok sa zvýšil na 566. Dôvodom tohto kroku bol vznik šiestich modelov v tréningovej fáze. Pre každého hovoriaceho boli vytvorené dva modely, každý model v jednej emócii.

UBM bol v oboch prípadoch natréningovaný pri použití GMM zmesi s 32, 64, 128, 256 a 512 komponentmi.

V procese extrakcie príznakov bolo získaných 19 a 22 mel frekvenčných kepstrálnych koeficientov (Mel Frequency Cepstral Coefficients - MFCC). Všetky koeficienty boli extrahované s použitím Hammingovho okna s dĺžkou 25 ms a posunutím 10 ms. V oboch prípadoch bola k počtu MFCC pridaná logaritmická energia, delta a akceleračné koeficienty. V druhom kroku sa k tomu istému počtu koeficientov pridali ešte terciárne regresné koeficienty. V prípade pridania prvých a druhých regresných koeficientov k počtu 19 MFCC bol výsledkom extrakcie vektor o veľkosti 60 koeficientov na rámec, v prípade 22 MFCC bol rozmer dimenzie extrahovaného vektora 69. V prípade pridania terciárnych regresných koeficientov k 19 MFCC, energii, delta a akceleračným koeficientom, výsledkom extrakcie bol vektor s dimenziou 80. Výsledný vektor o veľkosti 92 koeficientov na rámec vznikol, ak sa terciárne regresné koeficienty pridali k 22 MFCC, energii, delta a akceleračným koeficientom.

V procese odstránenia ticha boli energetické koeficienty normalizované v porovnaní s nulovým priemerom a varianciou. Následne boli vzniknuté príznaky filtrované na základe energie reči. Vektory s nižšou energiou akú mal porovnávací prah, boli odstránené. Na záver boli rámce s najvyššou energiou normalizované podľa pravidla priemeru a variancie.

V oboch prípadoch experimentu bola ako klasifikačná metóda použitá Mahalanobisova klasifikácia s EFR normalizáciou a CSS skórovanie s WCCN normalizáciou. Výsledky získané

pri použití Mahalanobisovej metódy pre prvý experiment sú zobrazené v Tabuľka 1, pre druhý experiment v Tabuľka 2. Výsledky rozpoznávania hovoriaceho získané použitím CSS sú v Tabuľka 3 pre prvý experiment a v Tabuľka 4 pre druhý.

Tabuľka 1 Verifikácia hovoriaceho s použitím Mahalanobisovej klasifikačnej metódy – 19 MFCC

	Experiment 1		Experiment 2	
	<b>MFCC 19 E_D_A</b>	<b>MFCC 19 E_D_A_T</b>	<b>MFCC 19 E_D_A</b>	<b>MFCC 19 E_D_A_T</b>
Počet komponentov GMM zmesi	Hodnota rozpoznávania [%]			
32	51.51	51.51	51.51	51.51
64	51.51	51.51	53.03	51.51
128	53.03	54.54	54.54	54.54
256	54.54	54.54	53.03	53.03
512	54.54	50.00	53.03	53.03

Tabuľka 2 Verifikácia hovoriaceho s použitím Mahalanobisovej klasifikačnej metódy – 22 MFCC

	Experiment 1		Experiment 2	
	<b>MFCC 22 E_D_A</b>	<b>MFCC 22 E_D_A_T</b>	<b>MFCC 22 E_D_A</b>	<b>MFCC 22 E_D_A_T</b>
Počet komponentov GMM zmesi	Hodnota rozpoznávania [%]			
32	63.63	63.63	66.66	66.66
64	74.24	74.24	75.75	75.75
128	74.24	81.81	83.33	84.84
256	66.66	66.66	68.18	68.18
512	66.66	66.66	66.66	68.18

Tabuľka 3 Verifikácia hovoriaceho s použitím CSS klasifikačnej metódy – 19 MFCC

	Experiment 1		Experiment 2	
	<b>MFCC 19 E_D_A</b>	<b>MFCC 19 E_D_A_T</b>	<b>MFCC 19 E_D_A</b>	<b>MFCC 19 E_D_A_T</b>
Počet komponentov GMM zmesi	Hodnota rozpoznávania [%]			
32	48.48	50.00	50.00	50.00
64	50.00	50.00	51.51	50.00
128	50.00	51.51	51.51	51.51
256	51.51	51.51	50.00	51.51
512	53.03	54.54	54.54	54.54

Tabuľka 4 Verifikácia hovoriaceho s použitím CSS klasifikačnej metódy – 22 MFCC

	Experiment 1		Experiment 2	
	<b>MFCC 22 E_D_A</b>	<b>MFCC 22 E_D_A_T</b>	<b>MFCC 22 E_D_A</b>	<b>MFCC 22 E_D_A_T</b>
Počet komponentov GMM zmesi	Hodnota rozpoznávania [%]			
32	59.09	54.54	56.06	56.06
64	71.21	71.21	71.21	71.21
128	72.72	72.72	74.24	74.24
256	46.96	48.48	60.60	62.12
512	68.18	69.69	68.18	66.66

## VII. ZÁVER

Hlavnou myšlienkou tejto práce bolo zlepšenie systému rozpoznávania hovoriaceho použitím emočných nahrávok hovoriaceho pri tréningu modelov.

V prvom experimente boli na tréning modelu hovoriaceho použité nahrávky dvoch emócií z emočnej databázy SUS. Najlepšie výsledky pri použití klasifikačnej metódy Mahalanobis boli zistené pri použití 22 MFCC s pridaním primárnych, sekundárnych a terciárnych regresných koeficientov pri natréningu UBM 128 komponentmi GMM zmesi. Hodnota rozpoznávania hovoriaceho v tomto prípade dosiahla úroveň 81.81%. V prípade CSS klasifikácie najlepšou hodnotou rozpoznávania hovoriaceho bola hodnota 72.72% pri použití 22 MFCC, energie, delta a akceleračných koeficientov. Ten istý výsledok bol dosiahnutý, ak sa k spomínaným koeficientom pridali aj terciárne regresné koeficienty. Oba výsledky vznikli pri tréningu UBM 128 komponentmi GMM zmesi.

V druhom experimente bol model hovoriaceho tréningovaný nahrávkami jednej emócie z SUS databázy. Výsledkom tohto prístupu bola lepšia hodnota rozpoznávania hovoriaceho. Mahalanobisovým skórovaním pri tréningu UBM 128 komponentmi GMM zmesi bola dosiahnutá hodnota rozpoznávania 84.84%. Vstupné príznaky v tomto prípade boli tvorené 22 MFCC s energiou, delta, akceleračnými a terciárnymi regresnými koeficientmi. CSS skórovacia metóda pri tom istom spôsobe tréningu UBM mala za výsledok hodnotu rozpoznávania hovoriaceho 74.24% v prípade, ak sa použilo 22 MFCC s energiou, delta a akceleračnými koeficientmi. Ten istý výsledok bol dosiahnutý aj v prípade, ak sa k tomu istému počtu MFCC pridal terciárne koeficienty.

Experiment ukázal, že použitie vyššieho počtu MFCC zvyšuje výsledky rozpoznávania a pridanie tretieho regresného koeficientu má na zlepšenie rozpoznávania hovoriaceho nepatrný vplyv. Z výsledkov tohto výskumu vyplýva taktiež to, že tréning modelu subjektu z nahrávok jednej emócie má vplyv na zlepšenie výsledkov rozpoznávania tohto subjektu.

V ďalšej výskumnej práci by sme sa chceli venovať rozšíreniu SUS emočnej databázy. Taktiež by sme chceli na tréning a testovanie systému verifikácie hovoriaceho použiť rozšírený súbor nahrávok mužských, ako aj ženských hovoriacich. V procese tréningu UBM sa chceme sústrediť na použitie vyššieho počtu komponentov GMM zmesi, konkrétne 1024 a 2048 komponentov.

## POĎAKOVANIE

Výskum prezentovaný v tomto článku bol podporovaný Kompetenčným centrom znalostných technológií pre inovácie produkčných systémov v priemysle a službách (ITMS kód 26220220155) (50%) a projektom VEGA 1/0386/12 (50%).

## LITERATÚRA

- [1] M. Hric, M. Chmulík, R. Jarina, "Comparison of selected classification methods in automatic speaker identification", *Komunikácie (Communications)*, 13 (4), pp. 20-24, 2011.
- [2] D. A. Reynolds, "A gaussian mixture modeling approach to text-independent speaker identification", Ph.D. thesis, Georgia Institute of Technology, 1992.
- [3] R. Kuhn, et al. "Eigenvoices for speaker adaptation", *ICSLP*. Vol. 98, pp. 1774-1777, 1998.
- [4] D. Matrouf, N. Scheffer, B.G., J.F. Bonastre, "A straightforward and efficient implementation of the factor analysis model for speaker verification", *INTERSPEECH*, pp. 1242-1245, 2007.
- [5] P. Kenny, G. Boulianne, P. Ouellet, P. Dumouchel, "Joint factor analysis versus eigenchannels in speaker recognition", *Audio, Speech, and Language Processing*, IEEE Transactions on, pp. 1435-1447, 2007.
- [6] P. Kenny, P. Ouellet, N. Dehak, V. Gupta, P. Dumouchel, "A study of inter-speaker variability in speaker verification", *IEEE Transactions on Audio, Speech and Language Processing*, vol. 16, no. 5, pp. 980 – 988, 2008.
- [7] P. Vizslay, J. Juhár, M. Pleva, "Two-dimensional linear subspace learning based on discriminant analysis of speech", *Journal of Electrical and Electronics Engineering*, 5 (1), p. 273-276, 2012.
- [8] N. Dehak, R. Dehak, P. Kenny, N. Brummer, P. Ouellet, P. Dumouchel, "Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification", *INTERSPEECH*, Vol. 9, pp. 1559-1562, 2009.
- [9] N. Dehak, et al. "Cosine similarity scoring without score normalization techniques", *Proc. Odyssey Speaker and Language Recognition Workshop*, pp. 71-75, 2010.
- [10] N. Dehak, "Discriminative and generative approaches for long- and short-term speaker characteristics modeling: Application to speaker verification", Ph.D. thesis, École de Technologie Supérieure, Montreal, QC, Canada, 2009.
- [11] A. Larcher, et al. "ALIZE 3.0-open source toolkit for state-of-the-art speaker recognition", *INTERSPEECH*, pp. 1-5, 2013.
- [12] P.M. Bousquet, D. Matrouf, J. F. Bonastre, "Intersession Compensation and Scoring Methods in the i-vectors Space for Speaker Recognition", *INTERSPEECH*, pp. 485-488, 2011.
- [13] W. M. Campbell, et al. "SVM based speaker verification using a GMM supervector kernel and NAP variability compensation", *ICASSP*, Vol. 1, pp. I-I, 2006.



# Prototyp laboratórneho zariadenia pre sušenie a vytvrdzovanie hrubých vrstiev

*Slavomír Kardoš, Pavol Cabúk*

Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

slavomir.kardos@tuke.sk, pavol.cabuk@tuke.sk

**Abstrakt** — Predmetom príspevku je prototyp laboratórneho zariadenia pre sušenie sieťotlačových šablón a vytvrdzovanie vrstiev pre hrubovrstvové technológie a technológie dosiek plošných spojov (DPS). Účelom vývoja zariadenia je zjednodušenie a automatizácia čiastkových procesov v technológiách tvorby hrubovrstvových motívov, čo má dopad na urýchlenie procesu sušenia, zníženie energetických nárokov na prevádzku a v neposlednom rade odbremenenie realizátora procesu. Zariadenie je kompatibilné so štandardnými sieťotlačovými rámami do rozmeru 290 x 250 mm. Zariadenie zahŕňa aj možnosť sušenia a vytvrdzovania fotocitlivých vrstiev nanášaných na povrch dosky plošných spojov.

**KLúčové slová** — hrubovrstvové technológie, sieťotlač, fotolitografia, DPS

## I. SUŠENIE A VYTVRDZOVANIE V TECHNOLÓGIÁCH TVORBY HRUBÝCH VRSTIEV

Procesy sušenia a vytvrdzovania sú významnou časťou výrobného procesu s podstatným vplyvom na konečné vlastnosti hrubých vrstiev. Pre procesy využívajúce sieťotlačové operácie je používaných viacero hrubovrstvových systémov, ktoré vyžadujú rozdielne parametre sušenia a vytvrdzovania. Rovnako je tomu v technológiách výroby dosiek plošných spojov. V oboch aplikačných oblastiach je aktuálny problém kontaminácie prachovými časticami.

Počas sušenia sa natlačená vrstva zbaví látok schopných odparenia, ku ktorým nepatrí iba voda, ale zmesi rôznych látok organického pôvodu. Tieto látky obvykle pôsobia ako riedidlá, či spomaľovače.

V priebehu vytvrdzovania sa elimináciou vody alebo iných riedidiel iniciuje chemická reakcia vedúca k polymerizácii zložiek pojiva organického alebo anorganického pôvodu. V uvedených procesoch sa vyskytujú aj ďalšie procesy, ktoré sú hromadne nazývané ako termické procesy, umožňujúce adhéziu hrubých vrstiev. Zložky hrubých migrujú do substrátu bez prítomnosti ďalších chemických látok alebo vody. Iniciátorom pri tejto reakcii je teplo.

Sušenie a vytvrdzovanie sú čiastkovými procesmi fotolitografických procesov využívaných v mikroeletechnológiách, ako aj v polygrafickom priemysle. V sieťotlačových technológiách je využívaná filmová predloha, cez ktorú je exponovaný motív. Na sieťovinový materiál je aplikovaná emulzia tak, aby bola vytvorená rovnomerná vrstva po celej ploche sieťoviny. Emulzia je citlivá na UV svetlo, preto počas procesu sušenia emulzie je potrebné ju nevystavovať svetlu s touto zložkou, vrátane slnečného žiarenia. Komerčne dostupným svetlocitlivými materiálmi pre tieto technológie sú na Azocol Z1, Azocol Z175 Gold, Dirasol Supercoat 915 a ďalšie. [1-3] Na vytvorenie motívu je potrebné priložiť na vysušenú sieťovinu pozitívny motív a následne povrch exponovať zdrojom svetla vhodnej vlnovej dĺžky a intenzity. Neexponovaný povrch je následne oplachom odstránený z povrchu sieťoviny. Výsledkom je negatívna sieťotlačová predloha. V technológiách výroby dosiek plošných spojov je využívané svetlo špecifickej vlnovej dĺžky na prenos vzoru motívu z fotomasky na svetlocitlivý fotorezistívny materiál. Výhodou fotolitografie je možnosť precízneho prenosu motívu. Vlastnosťou tejto techniky je požiadavka na planárny základový substrát a čisté prostredie.

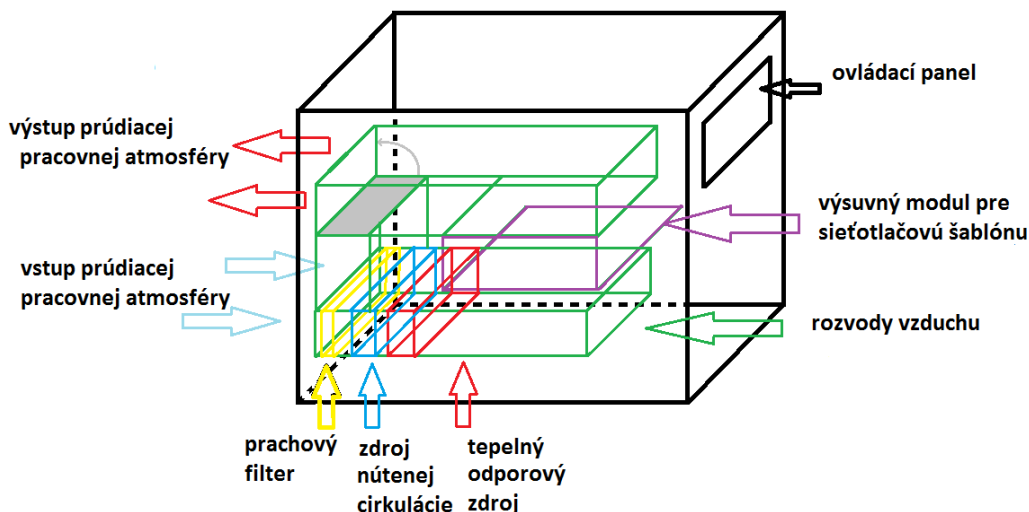
## II. LABORATÓRNE ZARIADENIE PRE SUŠENIE A VYTVRDZOVANIE HRUBÝCH VRSTIEV

Pre laboratórne a malokapacitné účely je dostupné množstvo komerčných zariadení pre sušenie a vytvrdzovanie rôznych štruktúr, pričom mnohé z nich umožňujú ventiláciu pracovnej atmosféry. Jedná sa o zariadenia pre všeobecné použitie, pričom pre sušenie sieťotlačových

motívov sú vhodné tie, u ktorých nemôže dôjsť ku korózii alebo inému poškodeniu zariadenia vplyvom stekajúcej vody. [4-5]

Pri návrhu boli zohľadnené požiadavky ako automatizácia procesu sušenia sieťotlačových šablón, nenáročná obsluha, optimalizácia spotreby elektrickej energie, redukcia kontaminácie pracovnej atmosféry a v neposlednom rade kompaktnosť.

V súlade s technologickými požiadavkami bolo laboratórne zariadenie pre sušenie sieťotlačových motívov navrhnuté z nasledovných častí: prachový filter, zdroj nútej cirkulácie pracovnej atmosféry, odporový tepelný zdroj, výsuvný modul pre sieťotlačovú šablónu, rozvody pracovnej atmosféry, uzatváracie klapky, ovládací panel a radiace obvody.



Obr. 1 Bloková konštrukcia teplotovzdušného zariadenia

Vo vstupnej časti zariadenia je umiestnený prachový filter pre zachytávanie prachových častíc, ktoré by mohli kontaminovať motív elektronického obvodu. Prachový filter je vytvorený z materiálu používaného na filtráciu prúdiacej atmosféry, napríklad na ochranu elektromotorov. Rám prachového filtra je vytvorený z laminátovej dosky. Za prachovým filtrom je v ceste tepelného zdroja umiestnená sústava ventilátorov, ktorých úlohou je nútená cirkulácia pracovnej atmosféry. Výkon tepelného zdroja bol zvolený s ohľadom na požiadavky rôznych procesov. Sústava rozvodov, ktoré distribuujú ohriatu atmosféru, je vytvorená zo sklolaminátových dosiek s obojstranne nalaminovanou medenou fóliou. Ich spájanie je realizované spájkovaním mäkkou spájkou. Povrch jednotlivých častí zariadenia je ošetrený antikoroziou vrstvou nanášanou v dvoch fázach. Základová fáza má zvýšenú adhéziu k povrchom ako sú medená fólia a pozinkovaný plech, ktorý je ako hlavný stavebný materiál konštrukcie realizovaného zariadenia. Prekrytá je syntetickou bielou farbou, poskytujúcou zvýšenou odolnosťou voči poveternostným podmienkam. Predný panel je opatrený striebornou metalízou a priesvitným lakom.

Výsuvný modul umožňuje vkladať a vyberať sieťotlačové šablóny. Je upravený tak, aby sa doň dali vkladať rôzne typy sieťotlačových rámov do rozmeru 290 x 250 mm.

### III. SNÍMANIE PARAMETROV PRACOVNEJ ATMOSFÉRY A RIADENIE SYSTÉMU

V ceste vystupujúcej pracovnej atmosféry sa nachádza jeden z integrovaných senzorov teploty a vlhkosti, ktorých úlohou je snímanie informácií o stave sušenia. Za zdrojom tepla je umiestnený druhý senzor teploty, ktorý slúži ako spätná väzba pre reguláciu teploty pracovnej atmosféry. Snímanie týchto pracovných parametrov poskytuje snímač teploty a absolútnej vlhkosti DHT22. Prostredníctvom digitálnej komunikácie sa snímač prepne z režimu nízkej spotreby do pracovného režimu a vyšle odozvu vo forme 40-bitovej informácie o relatívnej vlhkosti a teplote. [6-10]

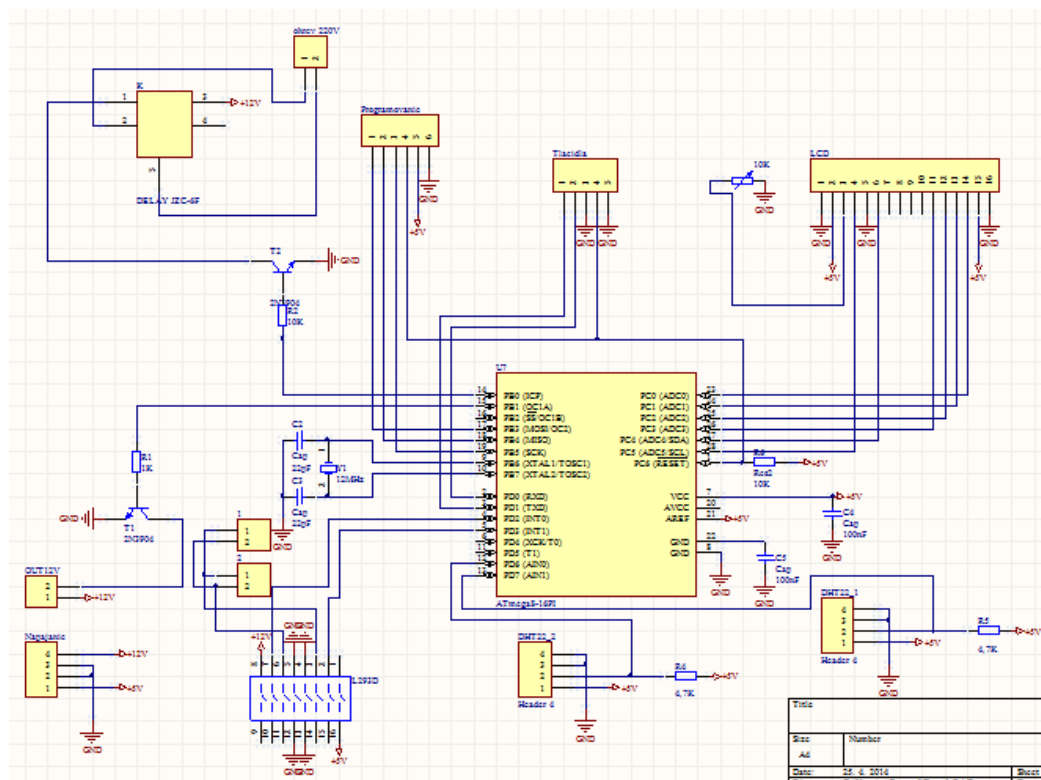
Zdrojom tepelnej energie je odporový vyhrievací element s núteným prietokom pracovnej atmosféry. Na zdroji sa nachádza tepelná ochrana vo forme termosínača, ktorý odpája napájanie zdroja, pokiaľ je prekročená kritická teplota. Regulácia zdroja je vykonávaná programovo spínaním zdroja tepla. Pre sušenie sieťotlačových motívov je potrebná teplota približne 35 až 40°C. V zariadení je možné dosiahnuť teplotu atmosféry do 70°C, čo je využívané pre vytvrdzovanie fotocitlivej emulzie pre maskovanie motívu DPS.

Pre riadenie systému je použitý mikrokontrolér ATmega8 z rady AVR od firmy Atmel. Vývoj softvérovej časti bol realizovaný v programe Atmel Studio 6.1. Program spracúva hodnoty od snímačov DHT22, zobrazuje ich na maticovom zobrazovači HD44780 s 16x2 zobrazovacími znakmi a zároveň ich používa na ovládanie procesov. [11-13] Klapky, ktoré zatvárajú a otvárajú prieduchy, sú ovládané elektromagnetickým vychyľovacím systémom, budený integrovaným H-mostikom L293D.

Užívateľské rozhranie sa nachádza v prednej časti zariadenia. Umiestnený je tam aj maticový displej a tlačidlá, pomocou ktorých je možné voliť pracovný mód.



Obr. 2 Predný panel s výsuvným modulom pre uloženie sieťotlačového rámu



Obr. 3 Obvodová schéma teplovzdušného zariadenia

#### IV. PROGRAMOVÉ PROFILY ZARIADENIA



Software laboratórneho zariadenia sa skladá z hlavnej časti a knižníc. Použité sú knižnice štandardných funkcií a navyše knižnice pre znakový displej HD44850 (16x2) a pre snímač teploty a vlhkosti DHT22. Aktualizáciu software je možné nahráť do zariadenia pomocou programovacieho konektora, ktorý sa nachádza na zadnej strane zariadenia.

Priebeh programu je sprevádzaný zobrazovaním informácií na displeji. Po zapnutí napájania a taktiež aj po stlačení tlačidla reset sa zobrazí na LCD paneli text Sušička a pod týmto textom verzia software. V úvodnej ponuke je volený režim sušenie sieťotlačového rámu a emulzie sieťoviny alebo režim sušenia fotocitlivého laku aplikovaného na povrch DPS. Po zvolení jedného z režimov zariadenie spustí cirkuláciu atmosféry a nastaví vnútorné klapky prieduchov pre zvolený režim. Zvyšnú časť procesu riadi mikrokontrolér, ktorý na základe času a snímania teploty a vlhkosti v zariadení ovláda zdroj tepla a klapky prieduchov.

Pri sušení sieťotlačových šablón sa otvoria klapky uzatvárania prívodu vonkajšej atmosféry a teplota sa reguluje približne na hodnote 35°C. Pri vytvrdzovaní fotorezistívneho laku sa vnútorné klapky najprv otvoria, aby sa odvetrali prchavé látky a následne sa uzatvoria, aby v zariadení cirkulovala ohriata atmosféra. Zariadenie sa vďaka tomu vyznačuje zníženou spotrebou elektrickej energie, pretože minimálne vyžaruje do okolia. Na zadanom mieste LCD displeja sa zobrazujú zostávajúce minúty a sekundy do uplynutia celkového priebehu programu. Pri znížení hodnoty vlhkosti pod kritickú úroveň je sušenie sieťotlačového rámu so sieťkou ukončené, nasleduje sekvencia dochladenia zdroja a zobrazenie stavu. Proces vytvrdzovania fotocitlivej vrstvy na DPS je ukončený iba podmienkou času. Po uplynutí dvadsiatich minút sa spustí dochladzovanie a následne je proces ukončený. Na pozadí oboch programov prebieha časovanie, ktoré zaisťuje, aby bol proces ukončený i napriek poruche snímačov. Po ukončení procesu je sieťovina alebo DPS pripravená na ďalšie spracovanie. Zariadenie je opätovne v prevádzke po stlačení tlačidla reset alebo po odpojení a opätovnom pripojení sieťového napájania.

#### V. ZÁVER

Realizované teplovzdušné zariadenie pre sušenie a vytvrdzovanie hrubovrstvových, predovšetkým fotocitlivých vrstiev je funkčným prototypom, ktorého funkcionality boli testované v štandardných laboratórnych podmienkach. Jeho prínosom je uľahčenie a automatizácia niektorých čiastkových procesov, a to sušenie sieťotlačových šablón po čistení, navrstvení fotocitlivej emulzie a jej vyvolaní. Zariadenie je rozšírené o funkciu sušenia fotorezistívneho laku nanoseného na povrch dosky plošných spojov.

	Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.
	Podporujeme výskumné aktivity na Slovensku. Projekt je spolufinancovaný zo zdrojov EÚ.

Poďakovanie patrí aj Ing. Erikovi Žofčinovi, ktorý sa v rámci svojej záverečnej práce podieľal na riešení problematiky.

#### LITERATÚRA

- [1] Azocol Z175 Gold, KIWO: Technické parametre, 2003. <<http://www.kiwo.de/en/products/produktseiten-kiwo/azocol-z-175-gold/>>.
- [2] Azocol Z1, KIWO: Technické parametre, 2003. <<http://www.cadisch.co.uk/asps/resources/pdf/51-2.pdf>>.
- [3] Dirasol supercoat 915, Sericom, Fujifilm, 2010. <<http://www.sericol.com/products/documents/SMP93DirasolSuperCoat915.pdf>>.
- [4] Programovateľná sušička Memmert, <<http://santiago.olx.cl/horno-esterilizador-memmert-sn-100-iid-487674622>>
- [5] Programovateľná sušička Binder FP 53, <[file:///C:/Users/Pou%C5%BE%C3%ADvate%C4%BE/Downloads/327\\_Ohrev.pdf](file:///C:/Users/Pou%C5%BE%C3%ADvate%C4%BE/Downloads/327_Ohrev.pdf)>
- [6] Radim Mutina, Meranie vlhkosti, stsproject, 2012. <<http://stsproject.net/?p=84>>
- [7] Senzor vlhkosti a teploty, D-robotic UK: DHT11. 2010. <<http://www.micro4you.com/files/sensor/DHT11.pdf>>.
- [8] Senzor vlhkosti a teploty, D-robotic UK: SHT1x humidity and temperature sensor. 2009. <<http://www.droboticsonline.com/index.php/arduino/sensor/sht1x-humidity-and-temperature-sensor.html>>.
- [9] Paul Regtien, Martin Halaj, Eva Kureková. Meranie teploty, <<http://www.kam.sjf.stuba.sk/katedra/publikacie/leonardo/ucebnica/08s.pdf>>
- [10] Snímanie vlhkosti. <<http://ap.urpi.fei.stuba.sk/sensorwiki/images/0/06/Vlhk.pdf>>
- [11] Mikroprocesory AVR, Atmel, ATmega128. 2011. <<http://www.droboticsonline.com/index.php/arduino/sensor/sht1x-humidity-and-temperature-sensor.html>>.
- [12] Ondrej Závodský, Programujeme AVR v jazyku C. <<http://svetelektro.com/clanky/kniha-programujeme-avr-v-jazyku-c-557.html>>
- [13] Robert Zahoranský, Mikroprocesor a mikrokontroler. <[http://zarsoft.dyndns.org/MyWeb/KAVS/VMO/Procesory\\_Atmel\\_%20AVR\\_Prednasky\\_ZR.pdf](http://zarsoft.dyndns.org/MyWeb/KAVS/VMO/Procesory_Atmel_%20AVR_Prednasky_ZR.pdf)>

# Vývoj výučbového zariadenia pre subtraktívnu prípravu DPS

*Slavomír KARDOŠ, Pavol CABÚK*

Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

slavomir.kardos@tuke.sk, pavol.cabuk@tuke.sk

**Abstrakt** — Predmetom príspevku je konštrukčný návrh a realizácia poloautomatického zariadenia pre subtraktívnu prípravu dosiek plošných spojov pre laboratórne a výučbové účely. Opísané sú jednotlivé konštrukčné a elektronické časti systému. Ovládanie zariadenia a priebeh technologického procesu je riadený komerčným mikrokontrolérom. Programové riadenie monitoruje stav hladiny pracovnej kvapaliny, jej teplotu, časuje podprocesy a ovláda jednotlivé výkonové periférie.

**Kľúčové slová** — subtraktívny postup, DPS, mikrokontrolér, fotolitografia

## I. ÚVOD

Subtraktívna technológia výroby vodivých motívov dosiek plošných spojov predstavuje základnú metódu ich výroby. Je to proces odstraňovania materiálu v špecifickej oblasti chemickým leptaním v kvapalnom médiu pre odstránenie vrstvy materiálu, ktorý nie je chránený organickou alebo kovovou maskou. Komerčné zariadenia pre túto technológiu sa v súvislosti s objemom výroby líšia konštrukciou a tvarom pracovnej komory, spôsobom, akým je pracovná kvapalina čerpaná, spôsobom jej ohrevu, typom použitých konštrukčných materiálov, spôsob ovládania a pripojiteľnosťou na systém tlakovej a odpadovej vody. V neposlednom rade je závažná otázka recyklácie pracovného média. [1-6]

## II. KONŠTRUKCIA ZARIADENIA PRE AUTOMATIZOVANÚ SUBTRAKTÍVNU PRÍPRAVU MOTÍVOV DOSIEK PLOŠNÝCH SPOJOV

Požiadavkami na systém zariadenia sú odolnosť voči agresívnemu prostrediu pracovných médií, priehľadnosť pracovnej komory kvôli možnosti sledovania procesu, automatizované napúšťanie a vypúšťanie pracovného média a oplachovej vody, jednoduchá manipulácia, pripojiteľnosť na systém tlakovej a odpadovej vody a maloobjemový charakter procesu.

Pri návrhu jednotlivých častí zariadenia bola venovaná pozornosť predovšetkým tvaru a vzhľadu zariadenia, použitej materiálovej a súčiastkovej základni vzhľadom na prítomnosť agresívnych látok, spôsobu naplňovania a vyprázdňovania pracovnej komory, spôsobu ohrievania pracovnej kvapaliny, snímaniu pracovných parametrov a spôsobu ovládania.

Pracovný okruh vyžaduje použitie materiálov odolných danému typu agresívneho prostredia. Sú to predovšetkým polymérne materiály, sklo a niektoré kovy. [7]

Konštrukciu pracovnej komory tvaru terária tvoria sklenené tabule a duté plastové profily vo funkcii prepážok. V spodnom plastovom profile sú zhotovené otvory o priemere 2,5 mm vzdialené od seba 15 mm, určené ako výstupné prebublávacie dýzy. Pracovný objem komory je 1,2 litra kvapaliny.

Bočné steny pracovnej komory sú vytvorené ako dvojkomorové tvoriac systém na princípe sífónu. Zlepenie bočných priečok bolo realizované univerzálnym silikónom, ktorý má vhodné tesniace a izolačné vlastnosti. Na prepojenia medzi kompresorom, výhrevným telesom, zásobníkom a komorou boli použité vinylové hadičky s vnútorným priemerom 6 a vonkajším 9 mm. Pre potreby zariadenia boli vyvinuté a vyrobené kapacitné snímače hladiny. Pre umiestnenie prvkov zariadenia bol skonštruovaný presúvateľný stojan. Na ovládanie zariadenia bol použitý vývojový modul s mikrokontrolérom ATmega 32.

Transport pracovnej kvapaliny z uzavretého zásobníka zabezpečuje kompresor. Zvolená bola metóda transportu kvapalného média pretlakom atmosféry nad jej povrchom. Pre nepriamy



ohrev pracovnej kvapaliny bolo použité výhrevné teleso s pomocným ventilátorom a umiestnené do PVC tuby o priemere 75 mm a dĺžke 250 mm. Ohrievacie teleso je umiestnené do cesty vzduchového kompresora smerom ku zásobníku. Následne po transporte pracovnej kvapaliny do pracovnej komory je táto prebublávaná vzduchom, a tým aj ohrievaná. [8, 9]

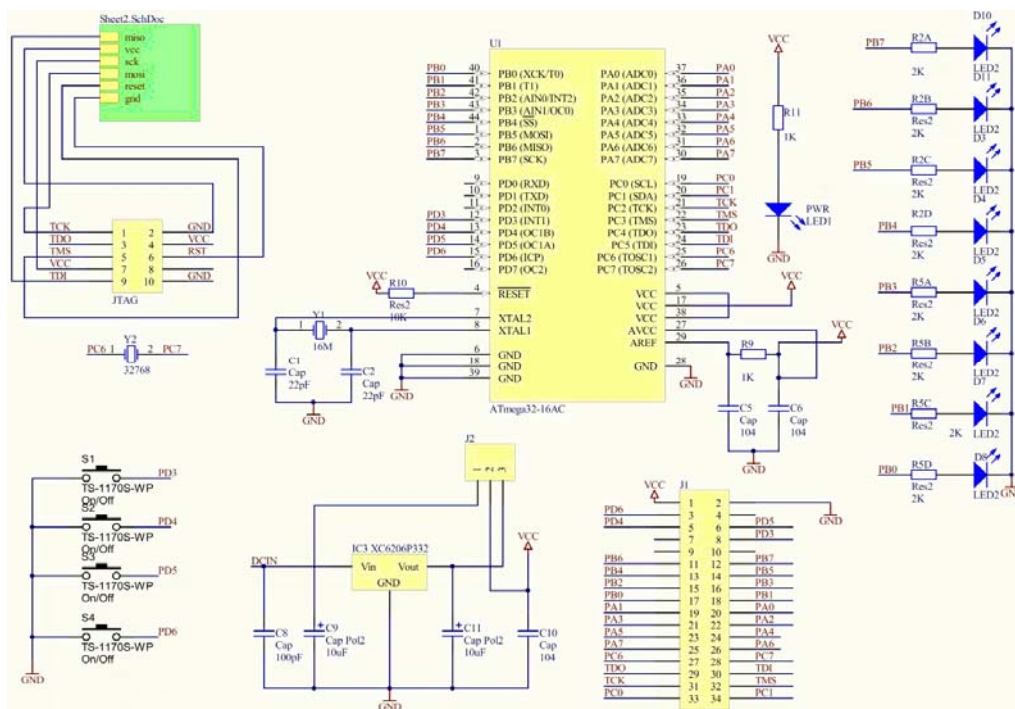
Oceľová procesná skriňa obsahuje spinaný zdroj s výstupnými jednosmernými napätiami 5 a 12 V, potrebnými pre napájanie elektronických častí zariadenia. Jej konštrukcia umožňuje efektívne rozmiestnenie jednotlivých prvkov.

Nosný stojan je riešený ako mobilný, aby ho bolo možné presúvať na požadované miesto a pripojiť na systém tlakovej a odpadovej vody. Stôl je vybavený dvoma oceľovými podstavami. Na spodnej podstave vo výške 150 mm je uložená procesná skriňa. Na vrchnej podstave vo výške 800 mm je umiestnená pracovná komora. Hladinové snímače sú umiestnené tak, aby bolo možné diagnostikovať hranice naplnenia komory.

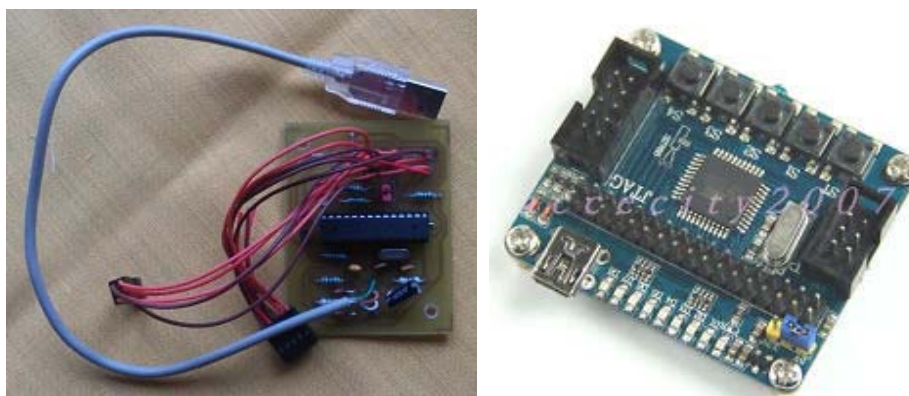
### III. ELEKTRONICKÁ A ELEKTRICKÁ VÝBAVA ZARIADENIA

Základnými funkčnými elektronickými blokmi zariadenia sú univerzálna doska s mikrokontrolérom ATmega32, USB programátor pripojený prostredníctvom ISP modulu s obvodi mi budenia periférií a napájací zdroj. Ovládacie prvky a riadiaca časť sú umiestnené na spoločnom prednom paneli.

Integrovaný programátor slúži pre uloženie užívateľského programu do pamäte mikrokontroléra a jeho čítanie cez rozhranie ISP. Po nahratí programu do pamäte mikrokontroléra slúži programátor už len pre napájanie mikrokontroléra. Na riadenie bol zvolený komerčný modul s riadiacim mikrokontrolérom ATmega32. [10]



Obr. 1 Schéma zapojenia riadiaceho modulu s mikrokontrolérom ATmega32



Obr. 2 Modul USB programátora (vľavo) a riadiaceho modulu s mikrokontrolérom ATmega32 (vpravo)

Pre zariadenie bol použitý bezolejový kompresor KNF NUEBERGER typu N811 KT.18 s asynchrónnym striedavým pohonom o výkone 50 W.

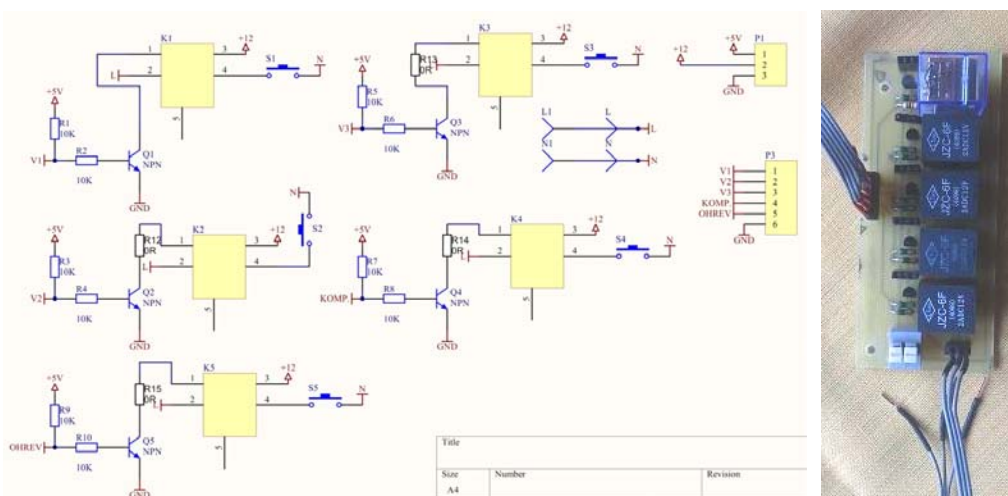
Pre snímanie minimálneho a maximálneho množstva pracovnej hladiny bol použitý kapacitný snímač hladiny na princípe šírkovej modulácie (PWM). Rozsah šírky impulzu bol nakalibrovaný na pracovný rozsah snímaného média. [11]

Snímanie aktuálnej teploty pracovného roztoku je riešené perličkovým termistorom NTCLE100E3103 - rezistorom s výrazným záporným teplotným koeficientom odporu. Zvolený termistor má pracovný rozsah teplôt od  $-40^{\circ}\text{C}$  do  $+125^{\circ}\text{C}$ . Nelinearitu vzťahu medzi odporom a teplotou v tejto aplikácii nie je potrebné osobitne riešiť. Citlivosť NTC snímača NTCLE100E3103 v pracovnom rozsahu teplôt je  $\pm 0,5$  až  $\pm 3\%/^{\circ}\text{C}$ . [12-16]



Obr. 3 Tubus s umiestneným odporovým telesom pre ohrev pracovnej kvapaliny

Obvody budenia periférií zariadenia sú určené pre ovládanie elektromagnetických ventilov, kompresora a ohrevu pracovnej kvapaliny. Ich riadenie je zabezpečené mikrokontrolérom ATmega32. Budenie relé je riešené spínacími tranzistorami.



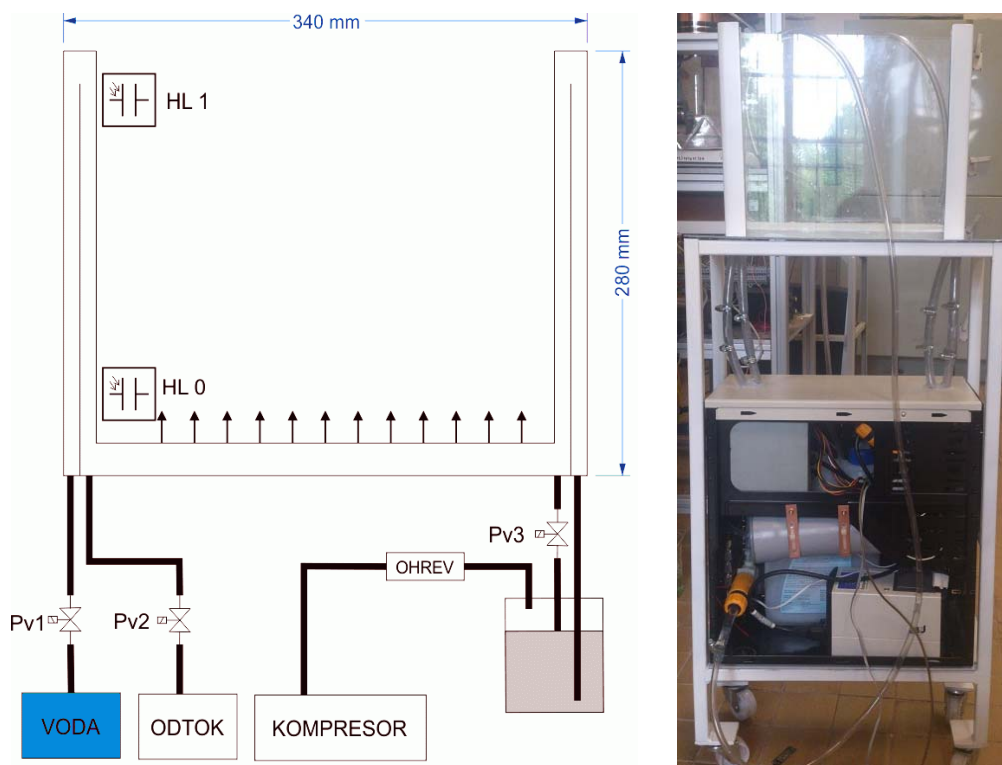
Obr. 4 Schéma zapojenia obvodu budenia periférií a osadený modul

#### IV. ŠTRUKTÚRA PROGRAMOVÉHO RIADENIA

Pripojením zariadenia na stabilizovaný zdroj program inicializuje nastavenia, rozsvieti zelenú LED na prednom paneli, nastaví, aby všetky elektromagnetické ventily boli uzavreté a všetky zariadenia vypnuté (kompresor a ohrev vzduchu). Program čaká v nekonečnej slučke a zisťuje stlačenie štartovacieho tlačidla. Po jeho stlačení zhasne zelená LED a rozsvieti sa červená LED.

Následne program zapne kompresor, aby sa prečerpala pracovná kvapalina do komory a zapne sa ohrev. Následne zisťuje, či už je prečerpaná. Ak áno, spusti sa časovač na 15 minút. Program je však možné prerušiť tlačidlom. Počas časovania program porovnáva hodnoty zo snímača teploty, a ak presiahne hodnota snímača teploty hodnotu  $50^{\circ}\text{C}$ , vypne ohrev. Po uplynutí časovania alebo stlačení tlačidla č. 2 program zhasne červenú LED, rozsvieti oranžovú, vypne kompresor a ohrev a otvorí elektromagnetický ventil V1. Ten bude otvorený, až pokiaľ sa nezmení hodnota na snímači HL0, čo signalizuje, že pracovná komora je prázdna. Nato sa zavrie ventil V1, a otvorí V2. Ostane otvorený dovtedy, pokiaľ program nezaznamená zmenu na snímači HL1. Táto zmena indikuje, že pracovná komora je naplnená oplachovou

vodou. To zabezpečí, že sa zavrie elektromagnetický ventil V2 a otvorí sa V3. Ten ostane otvorený, pokiaľ sa nezmení hodnota na snímači HL0 - pracovná komora je prázdna a zavrie sa elektromagnetický ventil V3. Otvorí sa ventil V2 a oplach sa opakuje. Po jeho ukončení je vypnutá oranžová LED a zapnutá zelená pre východiskový stav.





Obr. 5 Bloková schéma zariadenia a bočný pohľad na osadenú konštrukciu pred oživením (Pv1, Pv2, Pv3 – EMG ventily, HL 0 a HL 1 – snímače hladiny)

## V. ZÁVER

Účelom práce bolo navrhnuť a zrealizovať poloautomatické zariadenie, ktoré bude umožňovať subtraktívnou metódou vytvárať vodivé motívy dosiek plošných spojov. Návrh konštrukcie zariadenia bol nasledovaný výberom vhodnej súčiastkovej a materiálovej základne. Pracovná komora je maloobjemového charakteru, s možnosťou pozorovania procesu vo výučbovom procese. Riadenie zabezpečuje komerčný riadiaci modul s mikrokontrolérom ATmega32. Pre snímanie parametrov prebiehajúceho procesu boli navrhnuté a realizované diskretné snímače hladiny na báze PWM, ako aj obvod pre snímanie pracovnej teploty s termistorom typu NTC. Obvody budenia výkonových periférií pozostávajú zo spínacích relé budených tranzistorami. Mobilný oceľový rám obsahuje procesnú skriňu a pracovnú komoru, prepojenú vinylovými hadičkami. Predný panel procesnej skrine obsahuje riadiace a signalizačné prvky. V pracovnej komore je možné použiť rôzne komerčne používané pracovné médiá na báze  $\text{FeCl}_3$  a  $\text{CuCl}_2$  a potenciálne aj alkalické. Poloautomatické zariadenie pre subtraktívnu prípravu motívov DPS je vývojovým prototypom, ktorého niektoré funkcionality budú vylepšené.

## POĎAKOVANIE

 <p>Európska únia  <small>Tento projekt bol spolufinancovaný z prostriedkov Európskeho fondu regionálneho rozvoja.</small></p>	<p>Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.                  Podporujeme výskumné aktivity na Slovensku.                  Projekt je spolufinancovaný zo zdrojov EÚ.</p>
	

Poďakovanie patrí aj Ing. Miroslavovi Dluhošovi, ktorý sa v rámci svojej záverečnej práce podieľal na riešení problematiky.



LITERATÚRA

- [1] Kolektív autorov; Výroba plošných spojov, [Online], [cit. 26.10.2013.]. Dostupne na internete: <http://dpsk.sk/produkty/vyroba-plosnych-spojov/>.
- [2] Kolektív autorov; TechTerms.com, [Online], [cit. 10.11.2013.]. Dostupné na internete: <http://www.techterms.com/definition/pcb>.
- [3] Darling, R. B.; MicroFabrication: Wet Etching. [Online], [cit. 10.10.2013.]. <http://www.ee.washington.edu/research/microtech/cam/PROCESSES/PDF%20FILES/WetEtching.pdf>.
- [4] Kolektív autorov; PCB Basic, [Online], [cit. 26.11.2013], dostupne na internete: <https://learn.sparkfun.com/tutorials/pcb-basics>
- [5] Kolektív autorov; PCB Overview, [Online], [cit. 4.11.2013], dostupne na internete: <http://www.4pcb.com/pcb/>
- [6] Kolektív autorov; Alkalické leptanie DPS, [Online], [cit. 13.10.2013.], dostupne na internete: <http://svetelektro.com/clanky/alkalicke-leptanie-dps-532.html>
- [7] R. Hanfoug, A controlled wet etch process, 21.4.2011, [Online], [cit. 9.12.2013], dostupne na internete: <http://memslibrary.com/guest-articles/47-silicon-etching/11-a-controlled-wet-etch-process.html>
- [8] Hala Salman Hasan; Mathew Peet and H. K. D. H. Bhadeshia; Heat Transfer Experiments, [Online]. [cit. 18.12.2013.]. Dostupne na internete: [http://www.msm.cam.ac.uk/phase-trans/2007/HT/heat\\_transfer.html](http://www.msm.cam.ac.uk/phase-trans/2007/HT/heat_transfer.html)
- [9] Kolektív autorov; Direct vs indirect heating, [Online], [cit. 18.12.2013], dostupne na internete: <http://www.stalam.it/en/technologies/direct-vs-indirect-heating.html>
- [10] Atmel, ATmega32, [Online], [cit. 8.2.2014.], dostupne na internete: <http://www.atmel.com/devices/atmega32.aspx>
- [11] Preeti, Jain; Level sensors. [cit. 2.2.2014.], dostupne na internete: <http://www.engineersgarage.com/articles/what-is-level-sensor>
- [12] Kolektív autorov; Snímače teploty, [Online], [cit. 20.12.2013], dostupne na internete: <http://www.elektrotechnikaapc.tym.sk/obsah/teoria/clanky-priemyslovka/3.2%20snimace%20teploty.pdf>
- [13] Boghossian, Ardemis; Brown, James; Zak, Sara; 14.09.2006. Temperature Sensors. [cit. 2.2.2014.], dostupne na internete: <https://controls.engin.umich.edu/wiki/index.php/TemperatureSensors>
- [14] Kolektív autorov; Temperature Sensor – The Thermistor. [cit. 2.2.2014.], <http://www.facstaff.bucknell.edu/mastascu/elessonshtml/Sensors/TempR.html>
- [15] Kolektív autorov; NTC thermistor. [cit. 2.2.2014.], dostupne na internete: <http://www.resistorguide.com/ntc-thermistor/>
- [16] Thomas D. McGee; Principles and methods of temperature measurement. S. 203-236. [cit. 2.2.2014.], [http://books.google.sk/books?id=qfmS7g4JzjwC&lpq=PP1&pg=PA203&redir\\_esc=y#v=onepage&q=thermistor&f=false](http://books.google.sk/books?id=qfmS7g4JzjwC&lpq=PP1&pg=PA203&redir_esc=y#v=onepage&q=thermistor&f=false)
- [17] Kolektív autorov; Solenoid Valve Info, Solenoid Valve Basics, [Online], [cit. 18.1.2014.], dostupne na internete: <http://www.solenoid-valve-info.com/solenoid-valve-basics.html>

# Prieskumná Analýza Publikačnej Činnosti

*František BABIČ, Jakub Novák*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, jakub.novak.2@student.tuke.sk

**Abstrakt** — Článok popisuje proces prieskumnej analýzy dát reprezentujúcich publikačnú činnosť prostredníctvom vytvorenej aplikácie. Táto aplikácia umožňuje spracovanie, vizualizáciu a vyhodnotenie importovaných záznamov o publikačnej činnosti pomocou jednoduchých štatistických ukazovateľov alebo siete autorov. Takisto je možné v rámci aplikácie pridávať k záznamom konkrétne publikácie alebo dodatočne editovať ich vlastnosti. Na vytvorenie aplikácie nebol využitý žiaden redakčný systém, ale bola navrhnutá a implementovaná na mieru využitím programovacích a skriptovacích jazykov PHP, HTML, CSS, JavaScript, jQuery, MySQL a regulárnych výrazov pri importe dát.

**Kľúčové slová** — publikácie, grafy, sieť autorov

## I. ÚVOD

Publikačná činnosť predstavuje v akademickej sfére dôležitý zdroj informácií, nielen pri vyhodnocovaní úspešnosti jednotlivých pracovníkov, ale aj pri identifikácii kľúčových osôb, vzájomných väzieb alebo najčastejších foriem publikovania dosiahnutých vedeckých výsledkov alebo použitých metód. Existuje viacero databáz obsahujúcich záznamy o publikačných výstupoch, ale ponúkajú len obmedzené možnosti, ako si vytvoriť pokročilé pohľady na uložené dáta. Tento fakt motivoval návrh aplikácie, ktorej cieľom je ponúknuť používateľovi možnosť získať rôzne pohľady na dáta reprezentujúce publikačné výstupy, napr. možnosť vygenerovať sieť spolupracujúcich autorov, graficky zobrazíť počet publikácií daného autora v príslušnom roku alebo v príslušnom publikačnom médiu, atď.

## II. ZÁKLADNÉ PRINCÍPY APLIKÁCIE

Na začiatku bolo nutné identifikovať kľúčové požiadavky na aplikáciu, od ktorých sa odvíjal následne návrh funkčnosti, ktoré bude výsledná verzia aplikácie ponúkať, t.j.: možnosť nahráť export z elektronickej knižnice v presne definovanom formáte, extrahovanie potrebných informácií pre jednotlivé funkcie systému, sieť autorov, rôzne štatistiky a vizualizácie v čase, kategorizácia publikácií, príslušnosť k projektom (%), vyhľadávanie v publikáciách, možnosť zadať publikačné média.

Ako testovací zdroj dát bol použitý export publikačnej činnosti z Univerzitnej knižnice TUKE vo formáte rtf. Extrakcia informácií z tohto exportu tvorí kľúčovú časť celej aplikácie. Tento export je štruktúrovaný ako bežný textový dokument s dodržaním určitých štandardov, podľa ktorých systém tento dokument vytvára. Práve vďaka týmto pravidlám je možné spätne tento dokument načítať ako samostatné bloky textu s potrebnými informáciami a následne ich ďalej spracovať. Na extrakciu potrebných informácií boli použité presne zadefinované regulárne výrazy [1, 2]. Ako príklad uvádzame:

- `preg_match("/^[^:]+/u", $val, $autori)` – mená autorov,
- `preg_match("/:.(+?)\v/u", $val, $nazov)` – názov publikácie,
- `preg_match("/^\s+[-\s+([0-9]{4})\./u", $val, $rok)` – rok vydania publikácie.

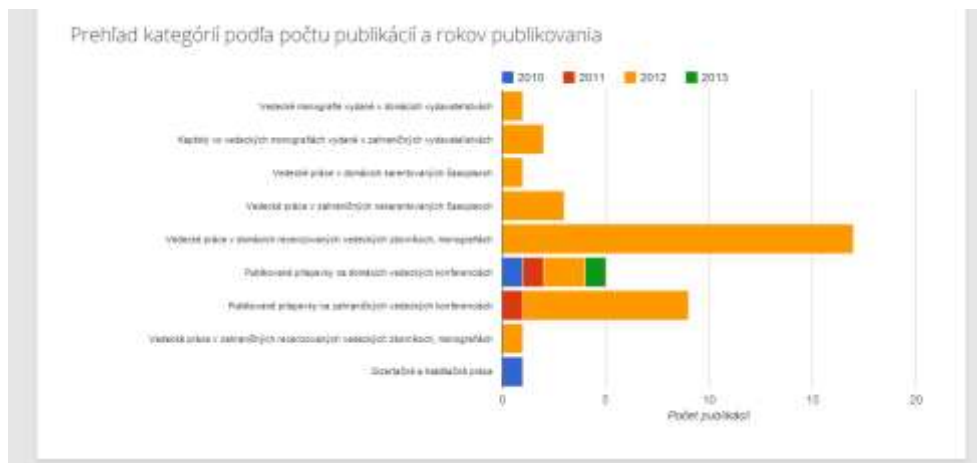
Pomocou týchto výrazov prebieha viacnásobné „strihanie“ textu, výsledkom ktorého sú relevantné úseky textu doplnené do formulára. Tento formulár je následne prezentovaný používateľovi, ktorý buď potvrdí extrahované informácie alebo ich ešte dodatočne upraví.

Sieť autorov poskytuje vizualizáciu prepojení medzi jednotlivými autormi importovaných publikácií, pričom na grafe môže byť zobrazená aj sila väzby medzi jednotlivými autormi interpretujúca množstvo spoločných publikácií, respektíve frekvenciu spolupráce. Silu väzby je možné v rámci vizuálnych prvkov znázorniť rôzne, napríklad farebnosťou (odtieň všeobecne použitej farby v grafe, stupňovaný podľa intenzity väzby) alebo hrúbkou čiary znázorňujúcej konkrétnu väzbu medzi dvomi autormi, pričom jej hrúbka by odzrkadľovala reálny počet publikácií nachádzajúcich sa v databáze systému. Na tento účel boli analyzované viaceré existujúce riešenia; ako najvhodnejší kandidát sa javila knižnica Arbor.js<sup>1</sup>, voľne dostupná pod MIT licenciou. Vykresľované grafy spĺňali požiadavky vizualizácie, zároveň ich bolo možné nastaviť do takej miery, aby bolo možné rozlíšiť jednotlivé elementy. Arbor.js zároveň disponuje responzívnym a interaktívnym rozhraním, pomocou ktorého bolo možné implementovať filtrovanie autorov a manuálne usporiadanie grafu. Výsledkom je interaktívny pavučinový graf, ktorý je možné usporiadať presunom jednotlivých elementov, vid'. obr.1. V hlavičke grafu je políčko, prostredníctvom ktorého je možné odfiltrovať autorov pre ich následné zobrazenie. Plochu grafu je možné ľubovoľne vertikálne zväčšiť alebo zmenšiť potiahnutím pravého dolného okraja plochy s malou ikonou troch šikmých čiar.



Obr. 1 Ukážka vizualizácie siete autorov pomocou interaktívneho pavučinového grafu.

Štatistiky sú v systéme zobrazené v dvoch formách: textovou a grafickou. Textová forma je reprezentovaná klasickým výpisom čísel alebo slov, ktoré sú buď priamo načítané z databázy alebo sú výsledkom sumarizačnej operácie, napr. počet publikácií konkrétneho autora. Grafická forma je postavená na rovnakom type informácií, na vykreslenie využíva knižnicu Google Charts<sup>2</sup>.

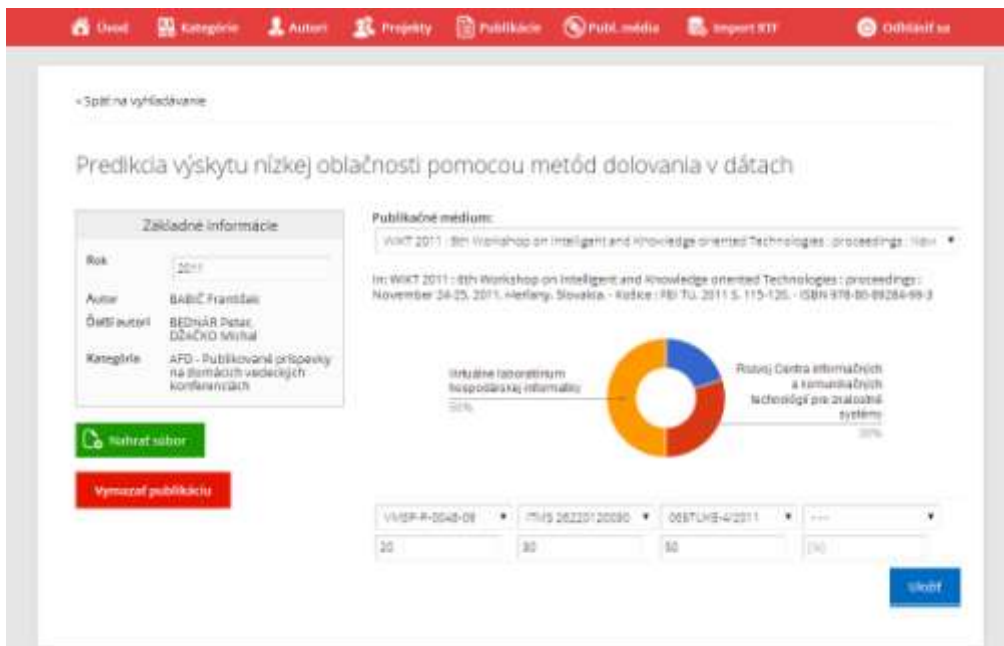


Obr. 2 Ukážka prehľadu publikačných kategórií podľa roku vydania publikácie.

<sup>1</sup> <http://arborjs.org>

<sup>2</sup> <https://developers.google.com/chart/>

Dôležitou funkciou, ktorej cieľom je evidovať informácie neobsiahnuté v aktuálnej verzii exportu, je možnosť zadať a editovať príslušnosť publikácie k projektom na základe poďakovania. Táto možnosť je súčasťou sekcie Publikácie, ktorá pre každý publikačný záznam obsahuje základné informácie, možnosť nahráť príslušný dokument, publikačné médium a príslušnosť k projektom.



Obr. 3 Ukážka detailu publikácie.

Prvý prototyp aplikácie bol testovaný viacerými pracovníkmi s cieľom overiť jeho vhodnosť a splnenie jednotlivých identifikovaných požiadaviek. Na základe tohto testovania boli identifikované a implementované nasledovné návrhy na zmeny: úprava privítania na úvodnej obrazovke; doplnenie sekcie s publikačnými médiami; viaceré úpravy regulárnych výrazov, ktoré špecifikovali rozdelenie obsahu súboru s exportom; zníženie dynamického usporiadania pri vykreslení siete autorov; prídanie identifikačnej položky *číslo* ku projektom; vytvorenie tlačidla pre zmazanie publikácie z databázy.

### III. ZÁVER

Finálna verzia aplikácie obsahuje všetky špecifikované funkčnosti a umožňuje používateľovi vytvoriť si rôzne pohľady na základe jeho požiadaviek alebo očakávaní. Najprv je však nutné importovať potrebné dáta reprezentujúce príslušné publikačné záznamy, čo by v prípade exportu vo formáte XML uľahčilo situáciu. Táto komplikácia však bola vyriešená využitím regulárnych výrazov, pomocou ktorých je textová reprezentácia publikačných záznamov rozparovaná podľa špecifikovaných podmienok. Dôležitou funkčnosťou je možnosť špecifikovať príslušnosť publikácie k projektom na základe informácií uvedených v poďakovaní. V budúcnosti by bolo vhodné popracovať na možnosti získať tieto informácie priamo pri importe daného dokumentu, prípadne automatizovať import dát s cieľom zabezpečiť aktuálnosť dát v databáze aplikácie.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] Stephen C. Kleene, "Representation of Events in Nerve Nets and Finite Automata", in Shannon, Claude E.; McCarthy, John. *Automata Studies*. Princeton University Press, 1956, s.3–42.
- [2] K. Thompson, K. "Programming Techniques: Regular expression search algorithm". *Communications of the ACM*, 11(6), 1968, s.419–422.

# Stereoskopické meranie vzdialenosti v automatickom inventarizačnom systéme dopravných značiek

<sup>1</sup>František RAKOCI, <sup>2</sup>Luboš OVSEŇÍK, <sup>3</sup>Ján TURÁN

<sup>1,2,3</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>ferorak@gmail.com, <sup>2</sup>lubos.ovsenik@tuke.sk, <sup>3</sup>jan.turan@tuke.sk

**Abstrakt** — Tento článok opisuje metódu merania vzdialenosti dopravných značiek od kamerového systému na základe stereoskopických obrazov. Stereoskopia je metóda používaná pre záznam a reprezentáciu 3D obrazov. Dokáže vytvoriť dojem priestoru použitím dvoch obrazov zaznamenaných vzájomne posunutými kamerami. Stereoskopické obrazy umožňujú výpočet vzdialenosti zaznamenaných objektov od kamerového systému. Tento výpočet je realizovaný na základe rozdielu v umiestnení vybraného objektu na oboch obrazoch. Výpočet vzdialenosti ovplyvňujú parametre, ako vzdialenosť medzi kamerami a ohniskové vzdialenosti oboch kamier. Táto metóda je alternatívou k laserovému meraniu vzdialeností. Implementáciou do automatického inventarizačného systému je možné zjednodušiť výpočet vzdialeností dopravných značiek od vozidla, znížiť obstarávacie náklady a v konečnom dôsledku zjednodušiť reálnu implementáciu inventarizačného systému v praxi.

**Kľúčové slová** — 3D obrazy, inventarizačný systém, meranie vzdialenosti, optický korelátor, stereoskopia, trigonometria

## I. INTRODUCTION

Čiastkovou úlohou automatického inventarizačného systému je zistenie presnej polohy snímanej dopravnej značky a porovnanie s očakávanou polohou. Nutnými informáciami pre tento výpočet sú GPS súradnice vozidla, pomocou ktorého sa inventarizácia vykonáva a taktiež jeho vzdialenosť od snímanej dopravnej značky. Prvotný návrh automatického inventarizačného systému zahŕňal pre tento účel použitie laserového diaľkometra. Metóda merania laserovým diaľkomerom patrí do skupiny aktívnych metód, ktorých predpokladom je vyslanie určitého signálu (laserový lúč, ultrazvukové vlnenie, atď) na meraný objekt a výpočet vzdialenosti na základe zaznamenaného odpovede v podobe signálu odrazeného od meraného objektu. Opačnou skupinou sú metódy, u ktorých je výpočet realizovaný iba na základe získanej informácie bez potreby explicitného budiaceho signálu. Ako nosné médium v tomto prípade slúžia napr. prirodzené svetelné lúče odrazené od objektu. Do tejto skupiny patrí stereoskopické meranie vzdialeností. Stereoskopickým meraním sa vzdialenosť objektu určuje na základe vzájomného umiestnenia meraného objektu na zodpovedajúcich obrazoch z dvoch vzájomne posunutých kamier. [7]

Prvá časť článku opisuje štruktúru stereoskopického kamerového systému. Sú tu znázornené základné prvky systému a ich geometrická konfigurácia. Následne sú vysvetlené vzťahy medzi jednotlivými veličinami a opísaný princíp výpočtu vzdialenosti meraného objektu od kamerového systému. V ďalšom je v článku opísaný algoritmus spracovania zosnímaných obrazov a získavania vstupných parametrov, potrebných pre výpočet vzdialenosti. Keďže použitie stereoskopického merania vyžaduje zmenu štruktúry vstupných periférií inventarizačného systému, je znázornený a vysvetlený rozdiel novej štruktúry oproti pôvodnému návrhu. V závere sú spomenuté výhody použitia stereoskopického merania ako alternatívy k laserovému meraniu, ale taktiež obmedzenia použitia a faktory ovplyvňujúce presnosť merania touto metódou.

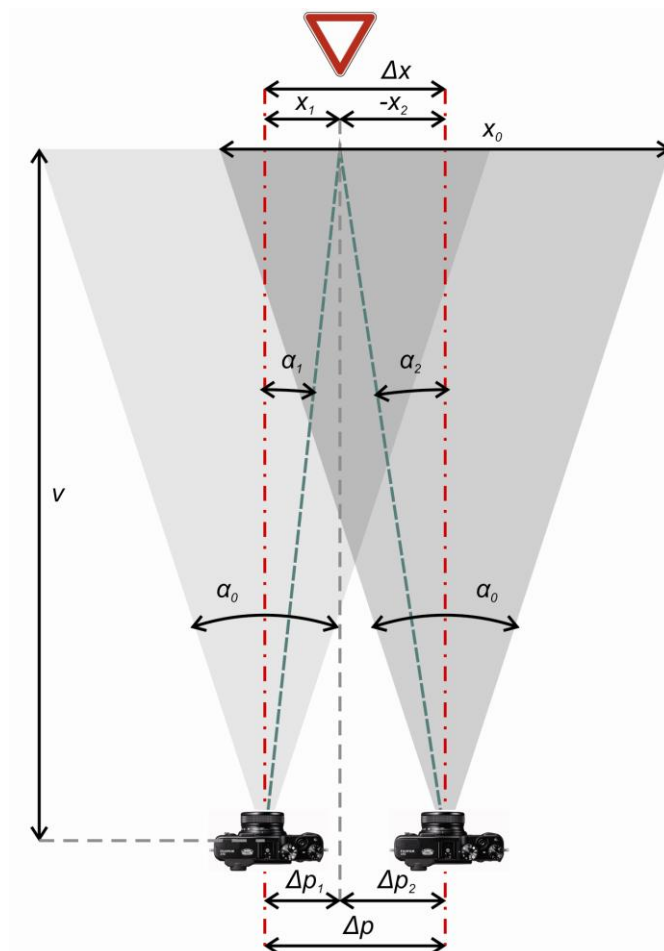
## II. STEREOSKOPICKÝ KAMEROVÝ SYSTÉM

Stereoskopický kamerový systém je konfigurácia optických zariadení, ktoré snímajú obraz z dvoch vzájomne posunutých pozícií. Môže pozostávať z jedného kamerového zariadenia s

dvomi šošovkami a prispôsobeným snímačom. Druhou, univerzálnejšou alternatívou je použitie dvoch nezávislých kamier, ktoré sú oproti sebe posunuté, čo má za následok súčasné snímanie dvojice obrazov s posunutou perspektívou.

Stereoskopia je metóda zaznamenávania a reprezentácie 3D obrazov. Dokáže vytvoriť dojem priestoru použitím dvoch obrazov s miernym vzájomným posunom. Stereoskopické obrazy môžu byť vytvorené dvomi kamerami. Predpokladom úspešnej reprodukcie je správna konfigurácia kamier a zaznamenanie scény obidvomi kamerami v rovnakom čase. Dôležitosť časovej synchronizácie je závislá od rýchlosti zmien v zaznamenávanej scéne. Pre statické obrazy je časová synchronizácia prakticky bezpredmetná. [2][3]

Nasledujúci obrázok znázorňuje geometrickú konfiguráciu kamier a zaznamenávanej scény.



Obr. 1 Štruktúra stereoskopického kamerového systému.

Ak je prvá kamera umiestnená na pozícii  $p_L$  a druhá na pozícii  $p_R$ , ich vzdialenosťou je

$$\Delta p = p_R - p_L \quad (1)$$

$\alpha_0$  predstavuje pozorovací uhol kamier. Výsledkom výpočtu je vzdialenosť  $v$  snímaného objektu od kamery. Vzdialenosť medzi kamerami môžeme vyjadriť ako súčet čiastkových vzdialeností kolmej projekcie objektu od jednotlivých osí pozorovacích uhlov kamier.

$$\Delta p = \Delta p_1 + \Delta p_2 \quad (2)$$

V prípade že osi pozorovacích uhlov oboch kamier sú paralelné, jej vzťah so vzdialenosťou meraného objektu je možné vyjadriť ako

$$\Delta p = \Delta p_1 - \Delta p_2 = v \cdot \tan \alpha_1 + v \cdot \tan \alpha_2 \quad (3)$$

Z toho vzdialenosť objektu od kamerového systému

$$v = \frac{\Delta p}{\tan \alpha_1 + \tan \alpha_2} \quad (4)$$

Ak  $\Delta x$  je posun obrazov z jednotlivých kamier vyjadrený v pixeloch a  $x_1, x_2$  sú pixelové vzdialenosti zosnímaného objektu od osi pozorovacieho uhla kamery, je možné určiť vzdialenosť meraného objektu na základe rozdielu umiestnenia tohto objektu na obrazoch získaných z oboch kamier. Platia nasledujúce vzťahy:

$$\frac{x_1}{\frac{x_0}{2}} = \frac{\tan \alpha_1}{\tan\left(\frac{\alpha_0}{2}\right)} \quad (5)$$

$$\frac{-x_2}{\frac{x_0}{2}} = \frac{\tan \alpha_2}{\tan\left(\frac{\alpha_0}{2}\right)} \quad (6)$$

Vzdialenosť objektu od kamerového systému je potom vypočítaná ako:

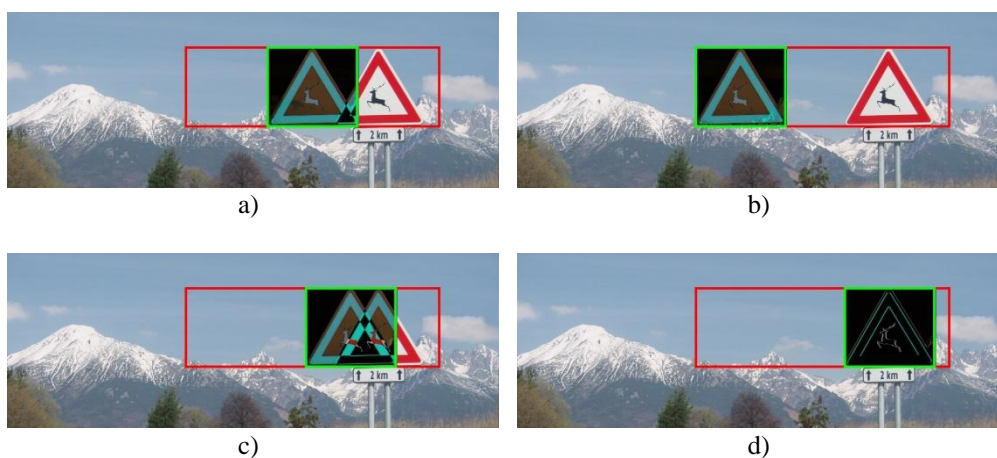
$$v = \frac{\Delta x}{2 \cdot \tan\left(\frac{\alpha_0}{2}\right)(x_2 - x_1)} \quad (7)$$

### III. ALGORITMUS ZÍSKAVANIA PARAMETROV

Pre určenie rozdielu v umiestnení meraného objektu medzi príslušnými obrazmi z oboch kamier je potrebné nájsť oblasti, ktoré v týchto obrazoch tomuto objektu zodpovedajú. V prvom kroku je potrebné vybrať túto oblasť v obraze z kamery umiestnenej vľavo. Príslušná oblasť je potom automaticky vyhľadávaná v obraze z kamery umiestnenej vpravo. Keďže poznáme geometrické rozloženie oboch kamier, vieme určiť, že objekt medzi týmito obrazmi je posunutý v horizontálnej osi. Vďaka tomu môžeme zdefinovať oblasť pre vyhľadávanie meraného objektu v obraze z kamery umiestnenej vpravo. Výška a vertikálna pozícia tejto oblasti sú zhodné s výškou a vertikálnou pozíciou oblasti reprezentujúcej vybraný objekt v ľavom obraze. Šírka prehľadávanej oblasti je z oboch strán rozšírená o okolie zodpovedajúce vzájomnému horizontálnemu odstupu použitých kamier.

Zodpovedajúce umiestnenie meraného objektu v pravom obraze je určené metódou najmenšieho rozdielu. Matica pixelov referenčnej oblasti sa porovnáva s maticami pixelov v prehľadávanej oblasti. Po otestovaní celého horizontálneho rozsahu sa vyberie umiestnenie, kde bol rozdiel minimálny. Nasledujúce obrázky demonštrujú postupnosť prehľadávania. Časť obrazu ohraničená červenou farbou je maticou možnej zhody s referenčnou maticou. Postupným posúvaním referenčnej matice (ohraničená zelenou farbou) je možné sledovať celkový rozdiel porovnávaných pixelov a následne vybrať umiestnenie, kde bol tento rozdiel minimálny. Vzdialenosť objektov medzi snímkami je určená rozdielom vzdialeností objektu od stredov zaznamenaných obrazov (osí pozorovacích uhlov).

Presnosť celého merania závisí od rozlíšenia obrazov, optického skreslenia a vzájomného umiestnenia kamier. [4][6]



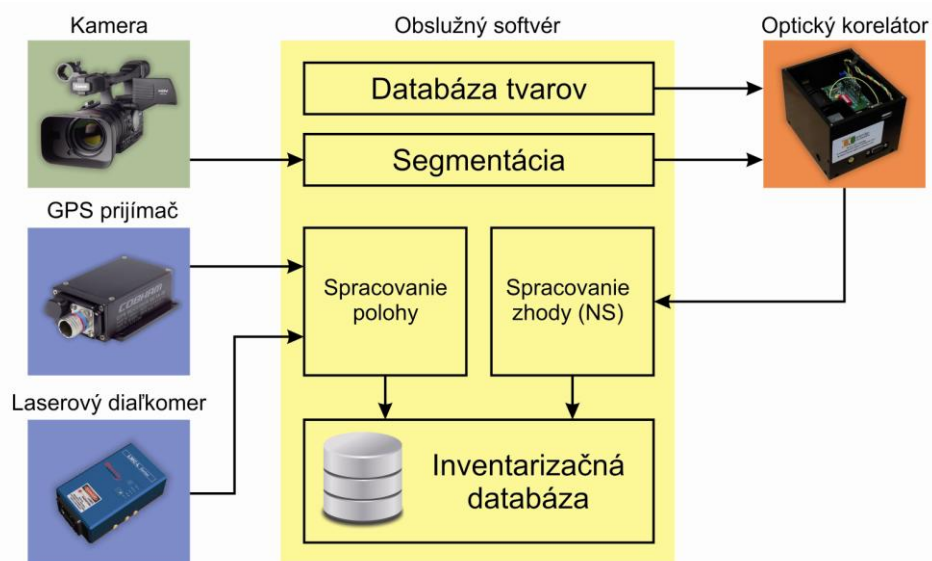
Obr. 2 Vyhľadávanie najmenšej chyby.

a) pôvodná pozícia objektu v referenčnom obraze, b) začiatok porovnávania v prehľadávanej oblasti, c) priblíženie k hľadanému umiestneniu, d) zhoda - minimálny rozdiel

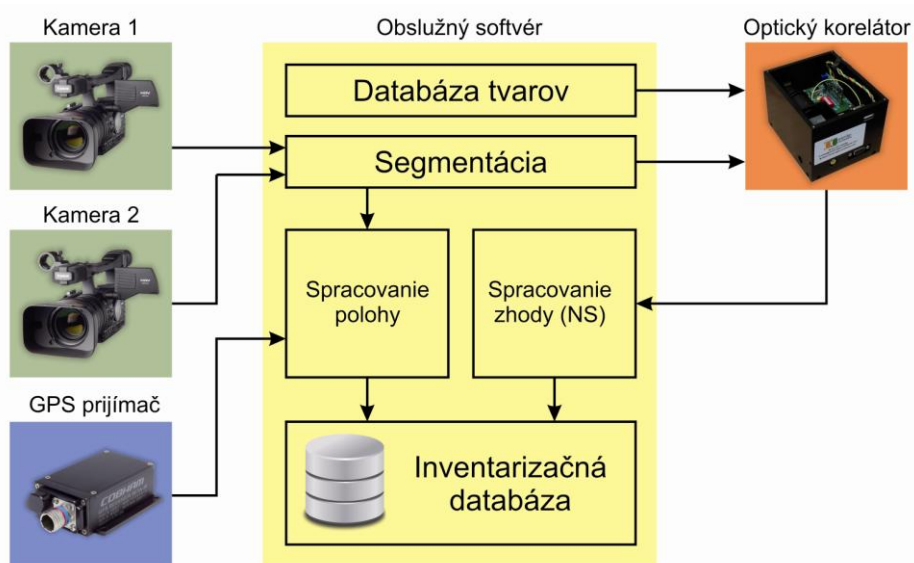


#### IV. IMPLEMENTÁCIA V AUTOMATICKOM INVENTARIZAČNOM SYSTÉME

Pôvodný návrh inventarizačného systému zahŕňal použitie laserového diaľkometra pre získanie vzdialenosti dopravnej značky od vozidla. Použitím stereoskopického merania je potrebné upraviť štruktúru vstupných periférií, kde sa namiesto laserového diaľkometra použije druhá, paralelná kamera.



Obr. 3 Pôvodný návrh inventarizačného systému s použitím laserového diaľkometra [5]



Obr. 4 Štruktúra inventarizačného systému so stereoskopickým meraním vzdialenosti

Implementáciou paralelnej kamery do štruktúry inventarizačného systému je možné realizovať meranie vzdialeností dopravných značiek od vozidla pomocou stereoskopických obrazov. Narozdiel od merania laserovým diaľkometerom, pri stereoskopickom meraní sú vstupné informácie získavané paralelne, s rovnakými parametrami bez potreby výraznej explicitnej synchronizácie. Obrazy z oboch kamier majú rovnaké rozlíšenie a snímajú scénu s rovnakým pozorovacím uhlom. Je preto relatívne jednoduché získať dostatočne presný výsledok pomocou relatívne jednoduchého algoritmu. Teoretickou možnosťou je určenie vzdialenosti pomocou optického korelátoru, kedy je možné na základe umiestnenia korelačných špičiek zistiť relatívnu polohu referenčného a porovnávaného obrazu. [1][8]

#### V. CONCLUSION

Stereoskopické meranie vzdialeností je založené na rozdielnom umiestnení sledovaného objektu medzi obrazmi získanými dvomi vzájomne posunutými kamerami. Vstupom pre



výpočet vzdialenosti sú dva paralelné obrazy. Zodpovedajúce umiestnenia konkrétnych objektov v týchto obrazoch sa určujú metódou najmenšieho rozdielu pri porovnávaní referenčnej oblasti z prvého obrazu s predpokladanou oblasťou z druhého obrazu. Implementáciou stereoskopického merania do automatického inventarizačného systému odpadá nutnosť synchronizácie medzi kamerou a laserovým diaľkomerom. Navyše, vďaka rovnakým parametrom vstupných obrazov je možné výrazne zjednodušiť celý proces merania. Nevýhodou stereoskopickej metódy je závislosť presnosti merania na rozlíšení obrazu. Náhradou laserového diaľkomera druhou, paralelnou kamerou je možné znížiť finančné nároky na realizáciu inventarizačného systému a tým pádom aj zjednodušiť reálnu implementáciu v praxi.

#### POĎAKOVANIE

Táto práca bola podporovaná Kultúrnou a edukačnou grantovou agentúrou (KEGA) Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky na základe projektu č. „063TUKE-4/2013 - Použitie diaľkovo ovládaného optického vláknového refraktometra vo výučbe“ a projektu č. „006TUKE-4/2014 - Experimentálny model PON siete TUKE a jeho použitie vo výučbe“. Táto práca je tiež výsledkom zmluvy „APVV-0025-12 - Predchádzanie vplyvu stochastických mechanizmov vo vysokorýchlostných plne optických sieťach“ podporovanej Agentúrou na podporu výskumu a vývoja (APVV) Slovenskej republiky.

#### REFERENCES

- [1] J. Goodman, *Introduction to Fourier Optics*, Greenwood Village: Roberts & Company Publishers, USA, 2004, ISBN 0-9747077-2-4
- [2] C. Holzmann, M. Hochgatterer, *Measuring Distance with Mobile Phones Using Single-Camera Stereo Vision*, 32<sup>nd</sup> International Conference on Distributed Computing Systems Workshops, 2012
- [3] A. Lorsakul, J. Suthakorn, *Traffic Traffic Sign Recognition for Intelligent Vehicle/Driver Assistance System Using Neural Network on OpenCV*, Proceedings of the 4th International Conference on Ubiquitous Robots and Ambient Intelligence, Postech, Piro, Korea, 2007, pp. 279 – 284
- [4] J. Mrovlje, D. Vrančič, *Distance measuring based on stereoscopic pictures*, 9<sup>th</sup> International PhD Workshop on Systems and Control: Young Generation Viewpoint, 1.-3. Október 2008, Izola, Slovenia.
- [5] F. Rakoci, *Automatický inventarizačný systém dopravných značiek s využitím optického korelátora*, Electrical Engineering and Informatics V, Proceeding of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, Slovakia, 2013, ISBN 978-80-553-1440-2
- [6] A. Yadav, T.B. Mohite-Patil, *Distance Measurement with Active & Passive Method*, International Journal of Computer Science and Network(IJCSN), Volume 1, Issue 4, August 2012, ISSN 2277-5420, pages 16-19
- [7] H. Zhaozheng, *Intelligent Road Sign Inventory (IRSI) with Image Recognition and Attribute Computation from Video Log*. Kyoto, Japan, 2011.
- [8] *Cambridge optical correlator*, published on <http://www.cambridgecorrelators.com/technology.html>

# Optimalizácia Internetového Predaja Pomocou Vhodnej Marketingovej Stratégie

*František BABIČ, Martin ANDERKO*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, martin.anderko@student.tuke.sk

**Abstrakt** — Tento článok je venovaný návrhu a implementácii vhodnej marketingovej kampane pre reálny internetový obchod s cieľom zvýšiť jeho návštevnosť a dosahovaný zisk. Analyzovaný obchod je postavený na platforme Open Cart, ktorá predstavuje jednoduchý, voľne dostupný e-commerce system, vhodný pre menšie až stredné firmy. Aplikovaná marketingová kampaň pozostávala z rôznych foriem propagácie a optimalizácie ako analýza návštevnosti, analýza databázy zákazníkov, SEO, PPC reklama, obsahový a sociálny marketing. Každá aplikovaná metóda bola následne vyhodnotená z pohľadu návštevnosti alebo generovaného zisku.

**Kľúčové slová** — e-shop, optimalizácia, analýza návštevnosti, SEO, PPC

## I. ÚVOD

Presadiť sa v internetovom podnikaní nie je v súčasnosti jednoduché. Denne vznikajú desiatky internetových obchodov, ktoré začínajú predávať tovar už existujúci na trhu a ich cenová politika nie je ničím výnimočná. Konkurencia vo väčšine oblastí je tak vysoká, že sa vstup na trh vôbec neoplatí. Mnohokrát je cieľom k úspechu práve výber správneho a efektívneho modelu internetového predaja, ktorý je cieľom tejto práce.

Optimalizácia a propagácia internetových obchodov zahŕňa množstvo pridružených oborov, ako je návrh webu, kódovanie, UX dizajn, copywriting, SEO, správa on-line reklamných kampaní, marketing na sociálnych sieťach a ďalšie. Každú činnosť pri propagácii je potrebné merať a analyzovať, aby sme dokázali vyhodnotiť, aký vplyv mali jednotlivé zmeny na výnosy a celkovú výkonnosť.

Analyzovaný obchod je postavený na platforme Open Cart, ktorá predstavuje jednoduchý voľne dostupný systém internetového predaja. Pred jeho prvým spustením v roku 2012 bola základná verzia rozšírená o viac ako 20 modulov a vylepšení zdrojového kódu, ktorými boli odstránené existujúce technické chyby a nedokonalosti. Po spustení bolo fungovanie e-shopu pravidelne monitorované prostredníctvom nástroja Google Analytics, prostredníctvom ktorého boli vytvorené čiastkové analýzy výkonnosti. Tieto analýzy ukázali nie veľmi efektívne využívanie finančných prostriedkov v reklamných systémoch, absenciu dôležitých zdrojov návštevnosti, nízke hodnoty konverzií a ďalšie problémy. Tieto dôvody následne viedli k návrhu a aplikácii modelu marketingovej stratégie, ktoré sú obsahom tohto článku.

Internetový marketing, PPC reklama, webová analytika, obsahový a sociálny marketing, SEO – pojmy, ktoré sa v posledných rokoch dostávajú do popredia stále viac. Webová analytika poskytuje podklady pre optimalizáciu a formovanie on-line, ale aj off-line stratégií. Je to činnosť zameraná na sledovanie vybraných ukazovateľov s cieľom zlepšiť výkonnosť webovej stránky (napr. tržby, konverzie a pod.). Analýza návštevníkov sleduje a snaží sa ponúknuť odpovede na otázky – ako a odkiaľ návštevníci na webovú stránku prišli, čo na webe robili, kým ho neopustili, prečo odišli a pod. Pri pokročilých webových analýzach sa sleduje množstvo metrik, ktoré sa navzájom kombinujú s rôznymi dimenziami (napr. vstupná stránka, poloha, trvanie relácie, výstupná stránka). Tieto dimenzie a metriky sa kombinujú za účelom čo najhlbšej analýzy dát (krížová segmentácia). Takto získané údaje majú veľký význam hlavne pre internetové obchody, ktoré vďaka nim môžu upraviť svoje marketingové a obchodné činnosti vo

svoj prospech [2].

## II. NÁVRH A APLIKÁCIA MARKETINGOVEJ STRATÉGIE

### A. Analýza Existujúceho Stav

Analyzovaný internetový obchod bol spustený v roku 2012 na základe identifikovaného potenciálu ponúkaného produktu s nízkou konkurenciou. Tento sezónny obchod je zameraný na predaj slnečných a čírych okuliarov. Cieľom bolo vytvorenie prehľadného a dobre organizovaného internetového obchodu, ktorý osloví široké spektrum zákazníkov; dosiahnutie vysokých pozícií vo vyhľadávačoch na rôzne kľúčové slová a samozrejme generovanie čo najvyššieho zisku.

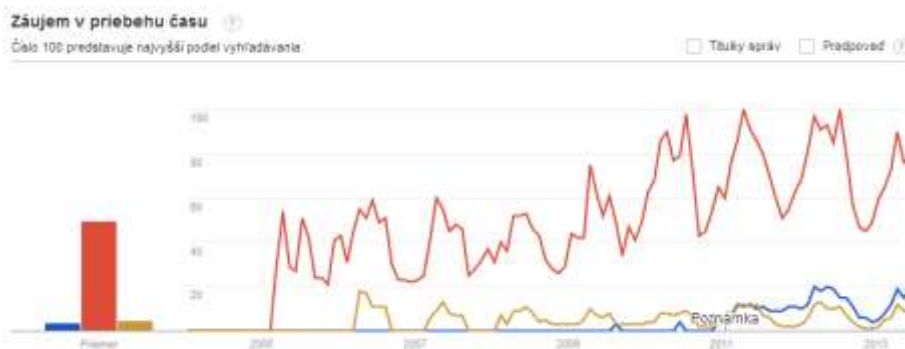
Pre účely vyhodnotenia aplikovanej marketingovej stratégie sme si stanovili nasledovné kritéria úspešnosti:

- zvýšiť podiel návštev z organického vyhľadávania o 20%,
- zvýšenie celkovej miery konverzie o 50%,
- nájsť nové zdroje návštevnosti,
- zvýšenie priamej návštevnosti o 20%.

### B. Analýza Kľúčových Slo

Kľúčové slovo predstavuje slovo alebo frázu, ktorú používateľ zadá do vyhľadávača. Cieľom webových stránok by malo byť práve zacieliť sa na relevantné kľúčové slová, aby následne boli zobrazené danému používateľovi na základe jeho dopytu. Výber vhodných kľúčových slov je dôležitou fázou optimalizácie, ktorej prechádza podrobná analýza, ktoré slová alebo frázy sú pre príslušný e-shop alebo webovú stránku relevantné, teda kľúčové. Väčšina podnikateľov a vývojárov sa začína zaoberať optimalizáciou pre vyhľadávače až po vytvorení webovej stránky, alebo e-shopu, čo je častá chyba.

Google Trends<sup>1</sup> predstavuje nástroj, ktorý poukazuje na trendy vo vyhľadávaní kľúčových slov, používateľské správanie vo vyhľadávaní k aktuálnym témam, čo ich zaujíma a o čo už návštevníci/používatelia strácajú záujem. Tento nástroj umožnil vytvoriť prvý pohľad na možné kľúčové slová pre analyzovaný e-shop. V tomto prípade sa ukázal kolísavý trend vo vyhľadávaní troch vybraných kľúčových slov: wayfarer, okuliare, slnečné okuliare, viď. Obr.1.



Obr. 1 Trendy vo vyhľadávaní 3 vybraných kľúčových slov.

Následne boli pomocou tohto nástroja jednotlivým vygenerovaným frázam pridelené váhy relevancie na stupnici od 0-100. Čím je číslo vyššie, tým vyššia je vyhľadávanosť (trend) daného slova v analyzovanom období, napr. wayfarer okuliare 100, okuliare 95, atď. Keďže kľúčová fráza „wayfarer okuliare“ dosiahla dva roky po sebe index 100, bola vybraná ako potenciálne vhodná na pomenovanie domény pre analyzovaný e-shop. Následne boli použité dva ďalšie nástroje na analýzu kľúčových slov (Keyword suggestion tool<sup>2</sup> a časť Google AdWords<sup>3</sup>), ktorých výsledky len potvrdili predchádzajúce zistenia a ponúkli zoznam slov, ktoré boli následne použité na aktualizáciu pôvodnej verzie e-shopu. Celková analýza kľúčových slov bola v tomto prípade použitá na vytvorenie pútavých popisov produktov a kategórií. Vhodné texty zvyšujú mieru konverzie, znižujú mieru odchodov a udržia návštevníka dlhšie na stránke. Z hľadiska optimalizácie kľúčových slov boli taktiež implementované recenzie produktov – kľúčové slová sa nachádzajú aj v textoch recenzií.

<sup>1</sup> <http://www.google.sk/trends/>

<sup>2</sup> <http://ubersuggest.org/>

<sup>3</sup> <https://support.google.com/adwords/answer/2999770?hl=sk>

Tento krok optimalizácie bol úspešný, na stránku eshopu sa po obsahovej zmene dostalo cez vyhľadávače pomerne viac návštevníkov. Pred optimalizáciou bolo stránka nájdená vo vyhľadávači Google pomocou 103 dopytov – po prevedení optimalizácie - 131 dopytov, čo malo vplyv aj na výnosy. Zmena konverzie bola veľmi badateľná. V analyzovanom období sa konverzia zmenila o 3.82%, čo je 252.8% zvýšenie (2.50% -> 6.32%).

### C. Analýza Zákazníkov

Predmetom analýzy v tomto kroku bola databáza zákazníkov, ktorá obsahovala viac ako 17 tisíc objednávok. Tieto dáta boli spracované vo forme dátovej kocky a analyzované prostredníctvom OLAP techník. Cieľovými metrikami boli: časové, geografické a demografické hľadisko realizovaných nákupov. Prvým zaujímavým zistením bolo, že najviac objednávok sa uskutočnilo v pondelok v doobedňajších a večerných hodinách. Na základe vysokej kúpnej sily a aktivity zákazníkov v tomto čase je dobré vtedy dostať na web čo najviac návštevníkov. Nedeľa a pondelok sú taktiež vhodnými dňami na spustenie on-line súťaží na sociálnych sieťach – v týchto dňoch je možné očakávať najväčší dosah a viralitu. Ako najsilnejší časový interval boli identifikované večerné hodiny od 19:00 do 22:00, čo je dôležitá informácia pre sociálny marketing a PPC reklamy (napr. Google Adwords, Facebook PPC), keďže je možné nastaviť častejšie zobrazovanie reklám v týchto časoch, zdieľať nové príspevky na sociálnych sieťach, rozosielať newslettere a podobne.

Z geografického hľadiska boli sledované obraty v jednotlivých krajoch Slovenska. Z analýzy vyplynulo, že Bratislavský, Žilinský a Prešovský kraj sú najsilnejšie z pohľadu uskutočnených objednávok. Bratislavský kraj má najväčšiu kúpnu silu, preto nás tieto hodnoty neprekvapili. Zaujímavý je Prešovský kraj, v ktorom obchod dosahoval vysoké obraty. Analyzovaný obchod má v tomto kraji odberné miesto, čo môže byť práve dôvodom tohto zistenia. Na základe týchto výsledkov je možné aplikovať dve opatrenia s cieľom zvýšiť obrat: prvým je otvorenie odberných miest v krajoch, kde je obrat najväčší – Bratislavský a Žilinský kraj. Druhou možnosťou je posilnenie slabších krajov otvorením odberných miest, keďže bolo zistené, že možnosť osobného odberu v Prešovskom kraji obchodu pomohlo.

Posledná analýza bola zameraná na vek zákazníkov, kde výsledné zistenia boli ďalej použité ako podklad pre tvorbu Facebook Ads<sup>4</sup>.

### D. SEO on-page

Skratka SEO vychádza z počiatočných písmen slov „search engine optimization“, čo sa prekladá ako optimalizácia pre vyhľadávače [1]. Keď hovoríme o vyhľadávačoch na Slovensku, máme na mysli hlavne fulltextový vyhľadávač Google. Pri on-page SEO ide hlavne o optimalizáciu technickej a obsahovej stránky. Na úvod boli realizované kontrola a úprava dôležitých HTML elementov (titulky, Meta Description a element IMG) a hyperlinkov. Pri kontrole a úprave titulcov bol použitý nástroj Xenu a Google Webmaster Tools. Veľmi dôležitý je taktiež element nadpisov - <h1> až <h6>; pri internetových obchodoch sa odporúča, aby sa hlavný nadpis h1 zhodoval s titulkom stránky. Pri používaní obrázkov na stránke – napr. produktové fotografie, ilustračné fotografie, je vhodné optimalizovať ich názov a alternatívny text. Alt-text vyhľadávač chápe ako popis obrázku a podľa tohto názvu ho indexuje. Takto je možné generovať návštevnosť aj z vyhľadávania v obrázkoch v Google. Pri textových odkazoch – napr. na podstránky, kategórie a pod. je vhodné odkazovať sa tzv. anchor textom, pomocou kľúčového slova. Týmto spôsobom vyhľadávač informujeme o tom, že cieľová stránka je relevantná pre dané kľúčové slovo.

Takisto bolo nutné vyriešiť duplicitu URL – s www a bez www. WWW.site.sk je subdoména domény site.sk, avšak väčšinou generuje rovnaký obsah, čo vyvoláva problém s duplicitnými stránkami. Rozloženie hodnotení a tzv. odkazovej štruktúry sa potom rozkladá po štruktúre stránky nerovnomerne a oslabujú sa celkové hodnotenia, čo je základný problém týchto duplicit. Riešením je presmerovanie verzie bez www na www verziu stránky (alebo naopak) pomocou http hlavičky 301. Pre stránky, ktoré bežia na servery Apache, je jednoduchým riešením súbor .htaccess, ktorý je uložený v koreňovom adresári hostingu.

Technika kanonických URL existuje už dosť dlho, avšak využíva ju stále málo webových stránok. Problém s duplicitným obsahom majú hlavne internetové obchody. Kanonická URL je v podstate preferovaná URL z viacerých URL, na ktorých je rovnaký alebo veľmi podobný obsah<sup>5</sup>. Vyhľadávač túto skutočnosť ľahko zistí a indexuje iba jednu URL, ktorú si sám vyberie,

<sup>4</sup> <https://www.facebook.com/advertising>

<sup>5</sup> <http://blog.pravidla-seo.sk/kedy-a-ako-urcovat-kanonicke-stranky/>

čo nie je veľmi vhodné. Riešením bolo na každú stránku e-shopu implementovať funkciu, ktorej úryvok zdrojového kódu vyhľadávaču napovie, ktorá adresa je nadržaná a ostatné vylúči z indexovania.

Robots.txt je súbor, ktorý určuje práva robotom na indexáciu určitého obsahu. Ak napríklad nechceme, aby roboti indexovali určitý adresár na webhostingu, vieme to určiť pomocou URL práve v tomto súbore. Sitemap.xml je súbor, ktorý obsahuje zoznam URL na danej doméne. Ide vlastne o mapu stránky, ktorá uľahčuje vyhľadávacím robotom indexáciu. Taktiež dáva odporúčania, ako často majú roboti danú stránku indexovať. Výstup súboru je vo formáte XML. Obe alternatívy boli v našom prípade povolené.

Open Graph Data sú informácie v zdrojovom kóde s účelom odporučiť sociálnej sieti Facebook, ktoré údaje o stránke ma zobrazit', akú fotografiu, alebo popis použiť pri zdieľaní. Ak tieto dáta nie sú pre URL definované, Facebook vyberá fotografie a výsek textu náhodne, ktorý následne pri zdieľaní zobrazuje. V analyzovanom e-shope boli implementované automaticky generované dáta, pre každú URL produktu. V týchto dátach je presne definovaná produktová fotografia, popis, URL a ďalšie údaje dôležité pri zdieľaní na tejto sociálnej sieti.

Výsledkom aplikácie vyššie popísaných opatrení bol vyšší počet vstupných stránok (viac ako o 100%), ktorý bol spôsobený najmä optimalizáciou štruktúry a architektúry jednotlivých častí e-shopu. Z pohľadu stanovených kritérií úspešnosti sa zvýšila návštevnosť o viac ako 130%, takisto miera konverzie o viac ako 100%.

### E. SEO off-page

Do off-page optimalizácie zaraďujeme všetky techniky a faktory, ktoré nie sú implementované na práve optimalizovanej webovej doméne. Ide hlavne o nákup spätných odkazov, registrácia do katalógov, vyhľadávačov tovarov, PR články, guest blogging a podobne. Jedná sa o tvorbu tzv. odkazového profilu, ktorý ovplyvňuje hodnotenie (ranking) stránky. Analyzovaný e-shop bol zaregistrovaný do kvalitných slovenských a českých katalógov; v prípade PR článkov išlo o platenú reklamu, pomocou ktorej bolo možné sa odkazovať na tento e-shop. Takisto bol e-shop zaregistrovaný na heureka.sk, kde priemerná konverzia presiahla hodnotu 7%, t.j. po prepočte boli náklady na jedného zákazníka približne 0.58€. Do tejto sekcie sme zaradili aj neplatenú reklamu prostredníctvom sociálnej siete Facebook. V tomto prípade boli využívané albumy s fotografiami produktov, súťaže na princípe zdieľania príspevku a užívateľskú súťaž. Výsledkom bola stála návštevnosť a výnosy, viď. Obr.3.



Obr. 3 Neplatená návštevnosť prostredníctvom sociálnej siete Facebook.

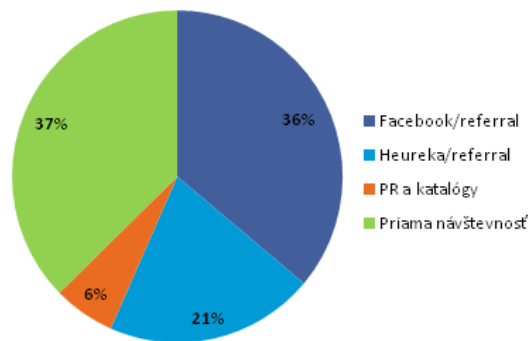
Pri tvorbe a sledovaní spätných odkazov bol použitý nástroj Majestic SEO<sup>6</sup>. V prehľadnej forme sú v ňom dostupné všetky spätné odkazy s ich hodnotením na základe rôznych ukazovateľov. Tvorba spätných odkazov je taktiež dlhodobá záležitosť. Nie je vhodné vytvoriť niekoľko spätných odkazov súčasne, keďže to vyhľadávač môže zaregistrovať ako hromadnú registráciu odkazov, z čoho plynie možná penalizácia. V odkazovom profile analyzovaného obchodu sa taktiež nachádzali nekvalitné spätné odkazy z katalógov s nízkym hodnotením, ktoré vznikli pri hromadnej registrácii v začiatkoch obchodu. Takéto odkazy sú nebezpečné, keďže môžu spôsobiť penalizáciu stránky, preto bola táto situácia vyriešená pomocou jednej z funkcií Google Webmaster Tools<sup>7</sup> (zabúdanie).

Celkovo je možné vyhodnotiť úspešnosť off-page optimalizácie pomocou výkonnosti jednotlivých zdrojov návštevnosti, viď. Obr. 4. Prekvapivo najvýkonnejšia bola priama návštevnosť, čo môže dokumentovať silu značky a trendovosť – ovplyvňuje ju celý

<sup>6</sup> <http://www.majesticseo.com/>

<sup>7</sup> [www.google.com/webmasters/tools/disavow-links-main](http://www.google.com/webmasters/tools/disavow-links-main)

marketingový mix vrátane platených reklamných kampaní. Heureka mala taktiež veľmi dobrú výkonnosť a vyznačovala sa vysokým ROI.



Obr. 4 Podiel jednotlivých zdrojov návštevnosti na výnosoch e-shopu.

#### F. PPC reklama

PPC alebo pay-per-click je spôsob reklamy, u ktorej sa platí za každé kliknutie na reklamu, teda za každého privedeného návštevníka. PPC reklama je veľmi dobrým doplnkom k SEO optimalizácii. Najpoužívanějšími PPC systémami Slovensku sú Google Adwords<sup>8</sup>, Facebook Ads a Etarget<sup>9</sup>. Na rozdiel od optimalizácie pre vyhľadávače nám PPC reklama prináša návštevníkov a výsledky prakticky ihneď po spustení kampane. Pri tomto e-shope boli použité všetky tri hore uvedené typy reklamy, ktorých výsledkom bol medziročný pokles nákladov na PPC o 13.92%, pričom výnosy z PPC stúpili o 36.3%, čo je veľmi dobrý výsledok.

### III. ZÁVER

Navrhnutý a realizovaný model optimalizácie internetového predaja dosiahol celkovo veľmi dobré výsledky. Návravnosť investovaných finančných prostriedkov počas jeho realizácie bola viac ako 500%. Počet vybavených objednávok za sledované obdobie presiahol 7 000. Miera konverzie elektronického obchodu z vyhľadávania celkovo vzrástla o viac ako 3.5%, čo malo za následok zvýšenie výnosov o 250% pri rovnakej návštevnosti – zvýšila sa efektivita internetového obchodu. Návštevnosť z vyhľadávania dôsledkom optimalizácie kontinuálne rastie a momentálne dosahuje priemerne 1500 - 2000 návštev mesačne pri miere konverzie približne 6%. Vo vyhľadávači Google tento elektronický obchod dosiahol popredné pozície na hlavné kľúčové slová. Do budúcnosti by bolo vhodné zvážiť využitie emailového marketingu, kontinuálne pokračovať v optimalizácii na rôzne kľúčové slová a long tail frázy.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] M. Kubiček, J. Linhart, „333 tipů a triků pro SEO“, Brno: Computer Press, a.s., 2010.
- [2] B. Clifton, „Google Analytics Podrobný průvodce webovými statistikami“, Brno, Computer Press, a. s., 2009.

<sup>8</sup> <https://adwords.google.sk>

<sup>9</sup> <http://www.etailer.sk/#1>

# Inventory System for Traffic Signs

<sup>1</sup>Ján RUŽBARSKÝ, <sup>2</sup>Ján TÓTH, <sup>3</sup>Ján TURÁN, <sup>4</sup>Luboš OVSEŇÍK

<sup>1,2,3,4</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>jan.ruzbarsky@tuke.sk, <sup>2</sup>jan.toth@tuke.sk, <sup>3</sup>jan.turan@tuke.sk, <sup>4</sup>lubos.ovsenik@tuke.sk

**Abstract** – At the age of technological advance, increasing of system performance and speed in every domain is needed do not forget on safety. More exactly on safety road users, which are the most sensitive member of traffic. Higher safety and auto-control of traffic signs is one of many topics investigated on Department of Electronics and Multimedia Communications in Kosice. In this paper is described proposal video-control system for vertical road traffic signs. It is not focused on detailed describe theoretical essentials, but demonstrates way of system functioning and basic goals. The project with name TSIS (Traffic Signs Inventory System) is still in progress. The main goal is to create an independent automatic system to check of traffic signs for different roads.

**Keywords** – Optical correlator, GPS, Inventory system, Traffic signs

## I. INTRODUCTION

Nowadays, high emphasis is put on safety in all sectors. The traffic sector is not an exception. The protection of passengers in a vehicle can be divided in to the passive and active. Passive protection includes safety belt or airbag, for instance. Active protection consists of systems that warn the driver of impending danger. It can be vehicles distance alarm, speeding alert, or warning of missing traffic sign.

The proposed system is designed to help solve last-mentioned problem. Brief description and history of traffic signs are in Chapter II. The shape and color of the signs are important factors in traffic sign recognition. Chapter III describes the proposal of traffic signs inventory system and hardware tools used in practical realization. That is an optical correlator (OC) designed at the University of Cambridge, a digital camera and a laptop. Chapter IV is devoted to the program TSIS (Traffic Signs Inventory System) to check the vertical road signs. Results of TSIS experiments are discussed in Chapter V. Conclusions of the proposed system are summarized in Chapter VI.

## II. TRAFFIC SIGNS

Traffic signs are the most important element of transport infrastructure today. They were evolved over time as the traffic grew. In the beginning there was only the need to indicate the directions at unclear crossroads, later also the distance to the certain objectives. Signs indicating the direction and distance, called milestones, were created already in the Roman Empire around 123 BC.

The largest shift in the road signs development was caused by development in cycling and then by the growth of automotive industry in the 19th century. First consolidation and standardization of the traffic signs was done successfully in 1908 at the meeting of the Permanent International Association of Road Congresses, PIARC, when 4 road signs were standardized. That was the birth of their present form. The most important milestones in the traffic signs development were years 1926 (Paris Convention), 1949 (Geneva Convention OSN, Protocol on Road Signs and Signals) and 1968 (Vienna Convention of European Economic Commission on Road Signs and Signals, OSN). The last mentioned, Vienna Convention from 1968 was the unifying standard that was accepted by most European and many other countries of the world. Overview of differences of some selected road signs is shown on Fig. 1.

In the picture we can see how particular signs differ, mainly in the background, icons and shapes. For a comparison we can take typical American diamond-shaped warning signs with yellow background and European ones in the shape of a triangle. It is worth to say that significant number of American signs is using only text form instead of generally known traffic pictograms [1].



Finland					
Poland					
Ireland					
Italy					
USA					

Fig. 1 Comparison of using traffic signs.



Fig. 2 Hardware implementation (camera PC correlator).

### III. SYSTEM DESIGN

The basic idea of the proposal hardware scheme of inventory system is shown in Fig. 2.

To create a traffic signs inventory it is necessary to make a video record of the selected part of the road. The GPS (Global Positioning System) coordinates corresponding to selected path must be obtained periodically in very small intervals and simultaneously with the video recording. In addition, we have to obtain also the GPS of every traffic signs along the route. All the data are processed by computer and for comparison and control of traffic signs the CC is used. To check and control the vertical traffic signs in off-line mode, video and GPS coordinates data of the road must be started. During the video playback the coordinates alter and at the same time they are compared with the data of traffic signs GPS coordinates. When a coincidence occurs between the positions obtained from the road and those of particular traffic signs, TSIS program takes a snapshot of the video and seeks for estimated traffic sign according to database. If the sign is recognized program shows the reference road sign following the traffic signs standard and the sign found in the video. Accurate and more detailed proposal of the whole concept process is shown in Fig. 3.

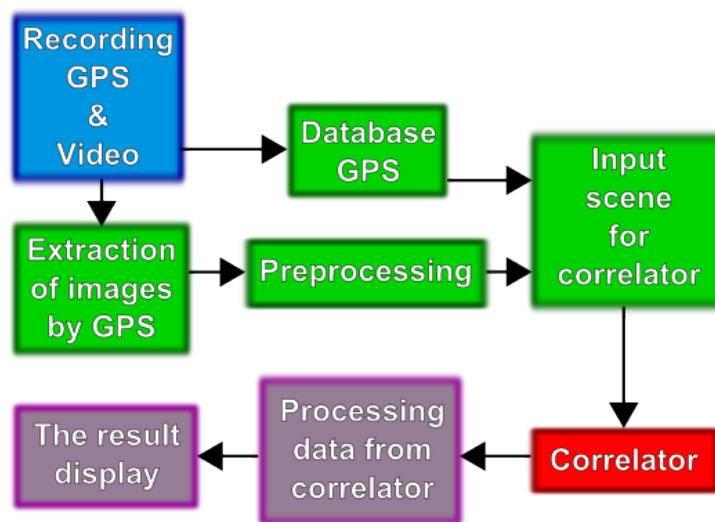


Fig. 3 Block diagram TSIS.



### A. An optical correlator

An OC identifies the content of the image by combining the incoming image with a reference image to determine the correlation intensity with the intensity of the light beam output. The OC's perform complex recognition faster than previously known digital techniques. They are able to process large quantity of data from the data stream and are useful in the detection, acquisition and classification of the information requested. They are used in many optical signal processing applications [2], [3].

Implementation of an optical interconnection in hybrid optoelectronic parallel computers is using technology of artificial neural network. Correlators are usually occurred in optical communication systems for signal detection. Data is transmitted over a light carrier, whose frequency and wave increase with the gradual development of technology. OC's implement the technology used spatially modulated light (SLM).

Since the implementation of the OC concept it is regarded as very useful especially in tracking techniques and military applications. Such applications are interesting because of their in-the-real-time image processing characteristics. With the implementation of this technology in to the commercial environment, the general view of this issue has changed. In spite of that applications are still more oriented on the industry. There is no doubt that the years of research invested in filters and algorithms will be reflected in the benefit of industrial as well as commercial applications. In our solution was implemented the OC created by the Department of Engineering at Cambridge University. This correlator uses Joint Transform Correlator (JTC) [4], [5], [6].

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

### B. Camera

The device contains two lenses. Front lens captures situation before the vehicle and the rear lens captures car interior. Wide angle of view provides enough space for retrospective image analysis. The front lens of the camera can still move vertically, so there is no problem to set the shooting area as required.

One of many views from vehicle is showed in Fig. 4. This view is from driver position but in reality it is recommended to situate the camera more on the right hand-side.

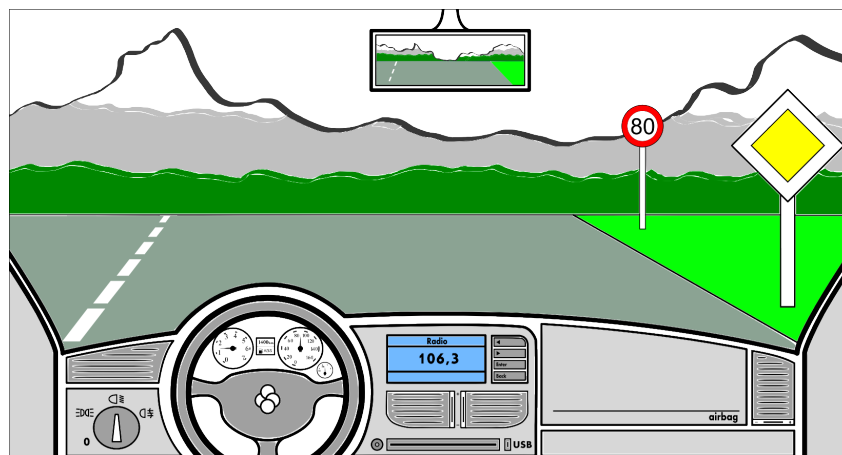


Fig. 4 View from vehicle.

On the picture above you can see one of many possibilities (Fig. 5.).

### C. Laptop

All processes associated with processing and video programming of TSIS were performed on laptop Lenovo Z580. Program for inventory of vertical traffic signs is written in Visual Studio 2008 in language C#. Language C# was selected because the software for correlator is also written in it.

## IV. TRAFFIC SIGN INVENTORY SYSTEM

Program for inventory of vertical traffic signs is written in Visual Studio 2008 in language C#. Language C# was selected because the software OC is also written in it.



Fig. 5 Camera in vehicle.

At the beginning it was necessary to select a section of the road, pass it, make a documentation on it - it means shoot a video, gather position signs on the route and record all the way to the GPS coordinates [7], [8].

After recording a video and collecting all the necessary data, the data were saved to the database via a graphical user interface. The proposed database contains reference traffic signs and to them GPS coordinates of selected road signs are assigned. While recording the video, GPS module in camera simultaneously created a KML text file where coordinates of the current location were stored in the specific time interval.

In the case of conformity of the coordinates of the KLM file and database program TSIS the image was taken from a video. Then an input scene for OC was created, which is composed of the reference signs according to the Decree of the Ministry of Interior of the Slovak Republic [9] and the processed scenes with brand search.

The user interface has three main windows through which data are processed. The first window - "Load information" is used to store information collected while driving a vehicle (Fig. 6).

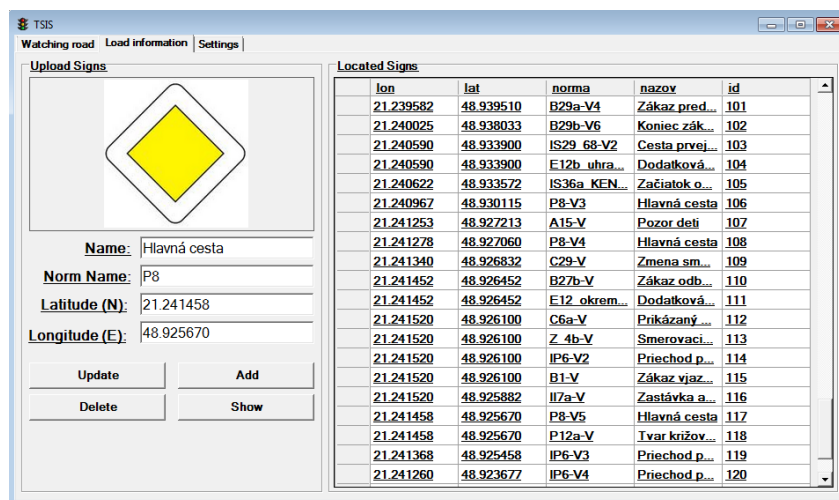


Fig. 6 Load information.

The second window "Watching road" - is surrounding for choosing the road and connecting attributes of KML file to video. In this window you check the control of signs and it is shows information about GPS position. This window is distributed in the three sectors. Sector Captured video displays video from a controlled way. Between Captured video and Signs is located Snapshot from video. In this block are showed moments when is in video captured traffic sign in GPS position accordance to information from database. In the sector Signs program displays reference traffic signs and extracted signs from the video. You can comparison original sign and sign from video. In the Fig. 7. you can see a preview of the functioning of TSIS.

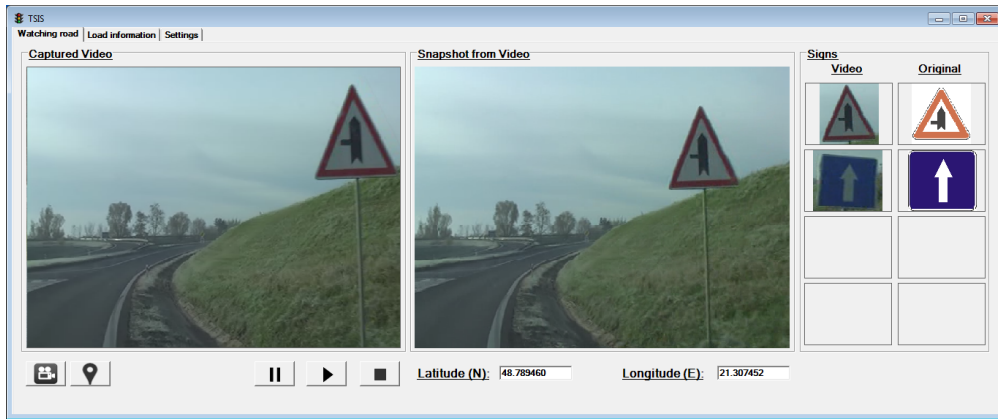


Fig. 7 Watching road.

The third window "Settings" allows to select the source data. It is more accurate way to coordinate database of traffic signs and reference signs (Fig. 8.). The reference signs are in common formats such as png or jpeg.

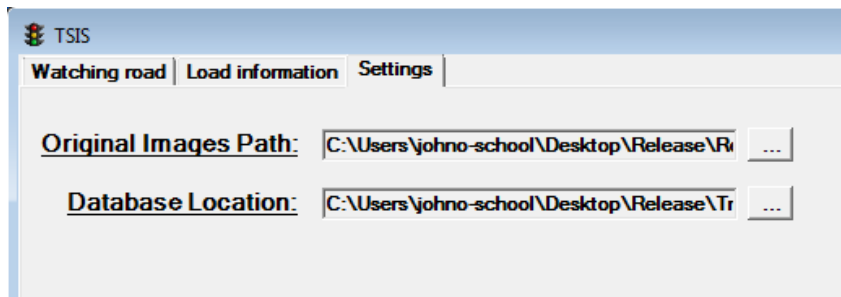


Fig. 8 Settings of TSIS.

## V. EXPERIMENTS

Path between the village Budimir and Kosice was chosen as experimental stretch. Selected path included 41 different traffic signs. The total number of signs was 73.



Fig. 9 Control of road.

Selected road section was controlled by TSIS. On a controlled path was not detected missing sign. System pointed to the situation as follows (Fig. 9.). TSIS in segment Signs showed all pairs signs (video-original).

## VI. CONCLUSION

During the inspection of vertical traffic signs this inventory system has not managed to uncover the missing sign. When is not detect any missing sign on control stretch of road it means that all signs are on their place. It demonstrated the practical application of the proposed system. In the future the system could be applied for example in vehicles of transport inspectorate, and inspectorate of road control. Checking traffic signs has a very important role, as part of prevention against accidents. The case of missing label can lead to catastrophic consequences, such as destruction of property, or death.

## ACKNOWLEDGMENT

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. 063TUKÉ-4/2013 - The Use of Remote Controlled Optical Fibre Refractometer in Teaching, and project no. 006TUKÉ-4/2014 - The Use of TUKÉ PON Experimental Model in Teaching. This work is also the result of the contract no. APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks supported by the Slovak Research and Development Agency.

## REFERENCES

- [1] S. Toth and E. Krsak, "Traffic Sign Recognition and Localization for Databases of Traffic Signs," *Acta Electrotechnica et Informatica*, vol. 11, no. 4, pp. 31–35, Jan. 2011.
- [2] A. Bergeron, "Optical correlator for industrial applications, quality control and target tracking," *Sensor Review*, vol. 20, no. 4, pp. 316–321, 2000.
- [3] T. Harasthy, L. Ovsenik, and J. Turan, "Current summary of the practical using of optical correlators," *Acta Electrotechnica et Informatica*, vol. 12, no. 4, pp. 30–38, 2012.
- [4] Official web page of cambridge correlator. 2013-09-30. [Online]. Available: <http://www.cambridgecorrelators.com/types.html>
- [5] M. Fifik, L. Ovsenik, J. Turan, and K. Fazekas, "Experiments with a Transform based Traffic Sign Recognition System," in *Proc. Of 17th International Conference on System, Signals and Image Processing (IWSSIP) 2010*, vol. 18, no. 1, pp. 227–230. Rio de Janeiro, Brazil, June 17–19, 2010.
- [6] M. Fifik, L. Ovsenik, and J. Turan, "Traffic Signs Recognition Experiments with Transform based Traffic Sign Recognition System," in *17th Symposium IMEKO TC 4, 3rd Symposium IMEKO TC 19 and 15th IWADC Workshop, Instrumentation for the ICT Era, International Institute of Informatics and Systemics*. Kosice, Slovakia, September 8–10, 2010.
- [7] T. Harasthy, L. Ovsenik, and J. Turan, "Optical correlator based Traffic Signs Recognition," in *19th International Conference on Systems, Signals and Image Processing (IWSSIP) 2012*, ISBN: 978-3-200-02328-4. Vienna, Austria, April 11–13, 2012.
- [8] T. Harasthy, L. Ovsenik, J. Turan, and K. Fazekas, "Traffic Signs Recognition with Using Optical Correlator," in *Proc. 18th International Conference on Systems, Signals and Image Processing (IWSSIP) 2011*, ISBN: 978-9958-9966-1-0. Sarajevo, Bosnia and Herzegovina, June 16–18, 2011.
- [9] Ministry of interior of the slovak republic, no 9/2009. 2013-09-30. [Online]. Available: [http://www.zbierka.sk/sk/vyhľadavanie?filter\\_sent=1&filter\\_predpis\\_aspi\\_id=9%2F2009&q=](http://www.zbierka.sk/sk/vyhľadavanie?filter_sent=1&filter_predpis_aspi_id=9%2F2009&q=)

# Mobilná Aplikácia Využívajúca Prvky Umelej Inteligencie

*František BABIČ, Peter MATIK*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, peter.matik@student.tuke.sk

**Abstrakt** — Správna reakcia v prípade prvej pomoci je dôležitým faktorom pre následné zlepšenie zdravotného stavu. Na tento účel existujú rôzne pomôcky; od Karty prvej pomoci, ktorá je súčasťou autolekárníčky; cez rôzne kurzy, prezentácie alebo tutoriály; až po mobilné aplikácie. Práve posledná alternatíva nebola na území Slovenskej republiky donedávna veľmi rozšírená, čo motivovalo vznik tejto práce. Vývoj v oblasti IT riešení sa však neustále posúva dopredu, čo zároveň prinieslo vznik podobného riešenia na konci roka 2013. To len dokumentuje situáciu, že kombinácia IKT a zdravotných služieb má pomerne veľký potenciál a na Slovensku poskytuje dostatočne široký priestor na realizáciu rôznych podporných softvérových riešení.

**Kľúčové slová** — mobilná aplikácia, prvá pomoc, expertný systém

## I. ÚVOD

Analýza vyhľadávaných výrazov vo vyhľadávači Google ukázala, že v celosvetovom meradle v posledných rokoch klesá záujem o prvú pomoc. Naopak, v priestore Strednej a Východnej Európy bol zaznamenaný výrazný vzostup, čo môže byť spôsobené zvýšenou dostupnosťou rôznych materiálov v elektronickej forme. Na tento trend plynule nadväzujú mobilné aplikácie, ktoré nachádzajú uplatnenie v rôznych oblastiach nielen spoločenského, ale aj hospodárskeho života. Nasadenie mobilných technológií práve v oblasti starostlivosti o zdravie dokáže výrazne zvýšiť komfort, zlepšiť zdravotnú starostlivosť a dokonca aj rýchlejšie pomôcť v situáciách, v ktorých to doteraz nebolo z technologického hľadiska možné. Ako už bolo spomenuté vyššie, konkurenčným riešením na Slovensku je mobilná aplikácia Prvá pomoc<sup>1</sup>, ktorá bola vyhlásená druhou najlepšou aplikáciou za rok 2013 porotou súťaže Android Code 2013. Táto aplikácia ponúka zásady prvej pomoci, priame volanie na tiesňovú linku, možnosť otestovať si svoje vedomosti, atď.

Ďalšou možnosťou ako zvýšiť pridanú hodnotu takýchto podporných softvérových riešení je pridanie interakcie s používateľom, napríklad formou umelej inteligencie. Nasadenie expertných systémov (ES) v oblasti medicínskej diagnostiky je už pomerne overenou myšlienkou, napr. jeden z prvých expertných systémov MYCIN slúžil na diagnostiku istých infekčných chorôb a následne navrhol vhodnú liečbu [1]. Typický model medicínskeho ES funguje na jednoduchom princípe pravidlovo-orientovaného systému, kde každé pravidlo pozostáva z podmienky AK, obsahujúcej symptómy a dôsledku POTOM, ktorým je konkrétna choroba alebo spôsob liečby. Najdôležitejšou časťou takéhoto ES je implementácia bázy znalostí na základe získaných znalostí od expertov alebo z relevantnej literatúry. Dôležité je takisto, aby takáto báza znalostí bola neustále aktualizovaná a testovaná na konzistenciu.

## II. NÁVRH A IMPLEMENTÁCIA APLIKÁCIE

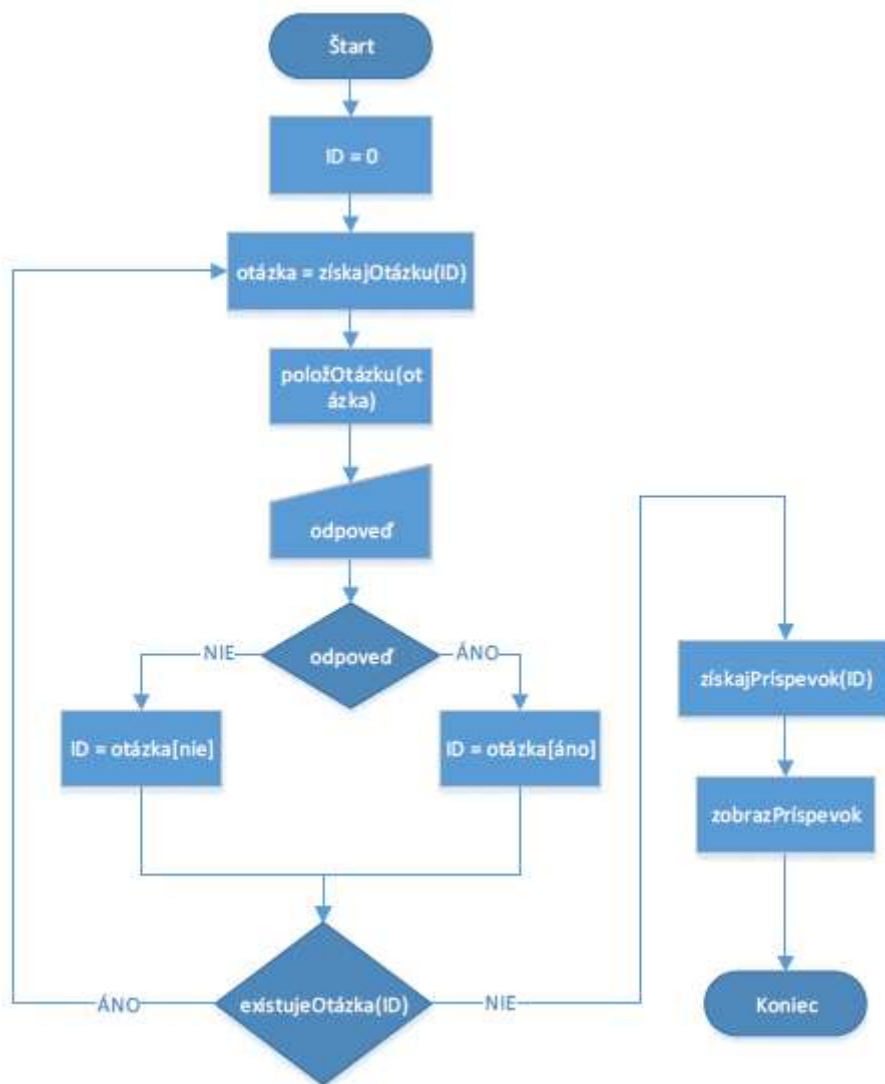
### A. Návrh Riešenia

Pri návrhu sme brali do úvahy viacero možných alternatív, aj vzhľadom na analýzu dostupných riešení a technológií; výsledkom je kombinácia webovej stránky a mobilnej aplikácie, ktoré ponúkajú totožný obsah. Z pohľadu používateľa boli identifikované nasledovné požiadavky: viacjazyčná podpora, moderný responzívny dizajn, prepojenie s mobilnou aplikáciou, využitie jednoduchého ES, prehľadný obsah a rýchla navigácia v ňom.

<sup>1</sup> <http://www.prvapomoc.sk/page/86/aplikacia/>

Z technologického pohľadu bol vybraný redakčný systém WordPress<sup>2</sup> ako vhodná alternatíva pre implementáciu webovej časti riešenia. Táto základná platforma bola následne upravená pomocou programovacích jazykov HTML, PHP, JavaScript a dostupných API tak, aby spĺňala stanovené používateľské požiadavky. Mobilná aplikácia bola vyvíjaná na platforme Android5 prostredníctvom SKD verzie API 15. Neskôr bola táto verzia rozšírená na verziu API 11, čo podľa viacerých prieskumov znamená, že výsledná aplikácia bude kompatibilná s väčšinou mobilných zariadení na trhu využívajúcich OS Android.

Z pohľadu ES bolo nutné vytvoriť inferenčný mechanizmus, ktorý na základe vstupov od používateľa bude vedieť extrahovať vhodné znalosti z bázy znalostí. Báza znalostí v tomto prípade bola tvorená príslušnými príspevkami z WordPressu. Algoritmus inferenčného mechanizmu je zobrazený na obr.1.



Obr. 1 Algoritmus inferenčného mechanizmu umožňujúci vyhodnotiť vstupy od používateľa.

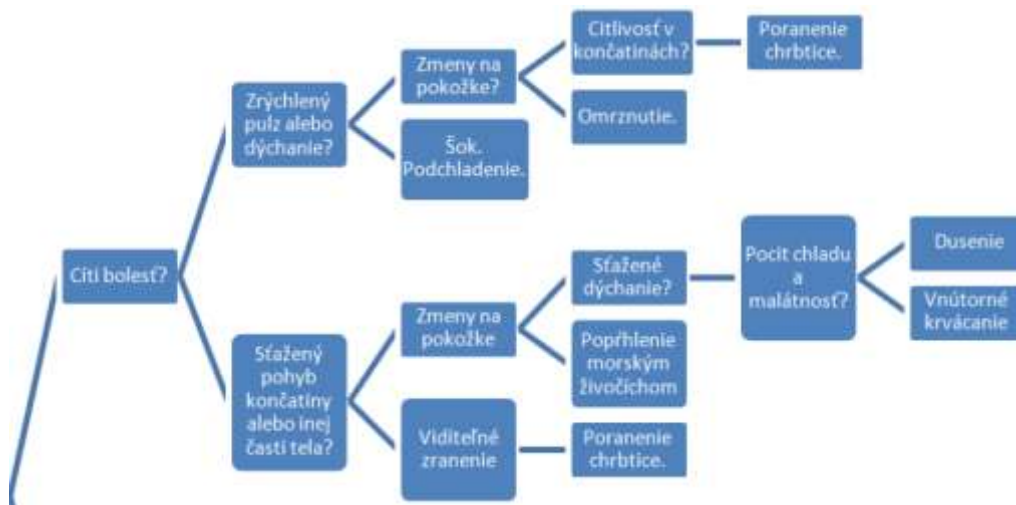
Po spustení algoritmu prebehne jeho inicializácia; nastaví sa ID prvej aktívnej otázky na hodnotu 0 (je dôležité, aby bola takto indexovaná prvá otázka inferenčnej siete už priamo v redakčnom systéme). V druhej fáze nasleduje nájdenie otázky v inferenčnej sieti na základe jej ID čísla. Po inicializácii nasleduje rozhodovací proces, ktorý prebieha v cykle a kladie používateľovi otázky. Algoritmus nie je automatizovaný a v tomto kroku čaká na vstup od používateľa, ktorým je kladná alebo negatívna odpoveď na otázku (používateľ má na výber len dve možnosti Áno a Nie). Na základe odpovede sa algoritmus rozhoduje, po ktorej vetve inferenčnej siete bude postupovať, teda ktorú ďalšiu otázku ponúkne používateľovi.

Podmienkou ukončenia cyklu je situácia, v ktorej algoritmus už nenašiel žiadnu novú otázku. V tomto prípade algoritmus predpokladá existujúce riešenie, ktorým je jeden alebo viacero

<sup>2</sup> <http://sk.wordpress.org/>



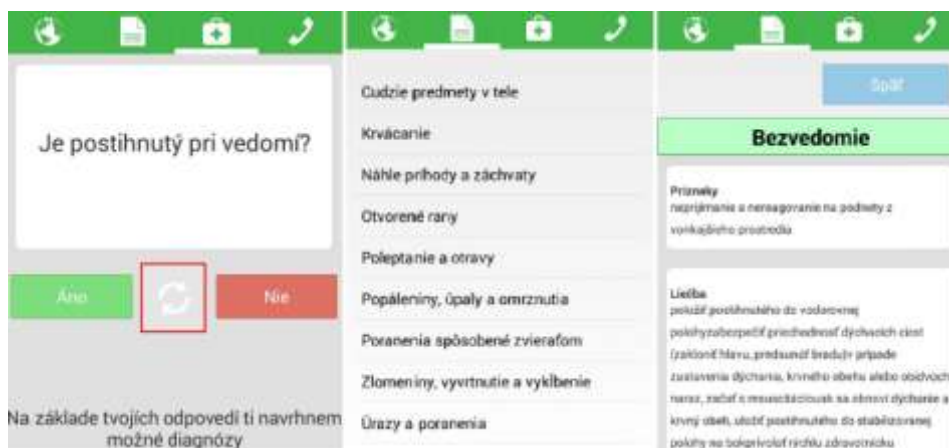
príspevkov identifikovaných týmto ID číslom. Do úvahy prichádza aj situácia, že riešenie situácie nie je jednoznačné a ES ponúkne viacero alternatív vyhovujúcich požiadavkám používateľa. To znamená, že príspevok môže byť identifikovaný viacerými ID číslami súčasne, prípadne dva príspevky môžu mať rovnaký identifikátor. Časť inferenčnej siete je zobrazená na obr.2.



Obr. 2 Ukážka inferenčnej siete, pomocou ktorej je možné diagnostikovať vhodný spôsob riešenia zdravotných problémov.

### B. Mobilná Aplikácia

Výsledná verzia mobilnej aplikácie je rozdelená do štyroch základných sekcií: úvodná obrazovka s jazykovými nastaveniami; zoznam kategórií; diagnostická pomôcka a zoznam telefónnych čísel. Na obr.3 je zobrazených niekoľko ukážok z aplikácie.



Obr. 3 Ukážky z mobilnej aplikácie.

Navigácia jednotlivých sekcií je zobrazená v hornej časti obrazovky v podobe záložiek. Používateľ sa medzi nimi prepína buď kliknutím na niektorú zo záložiek alebo tzv. ťahaním obrazoviek do strán. Úvodná obrazovka poskytuje používateľovi nástroje na správu obsahu a jazykov (momentálne je implementovaný slovenský a anglický). Okrem toho je na nej zobrazený povinný údaj o zodpovednosti za škody spôsobené používaním aplikácie a jej obsahu. Druhá sekcia predstavuje zoznam hlavných kategórií príspevkov; kliknutím na konkrétny príspevok sa obrazovka zmení a zobrazí sa obsah príslušného príspevku. Ten je v prehľadnej podobe rozdelený na príznaky, inštrukcie ošetrovania (liečba) a prípadne ďalšie poznámky. Sekcia diagnostiky ponúka používateľovi jednoduché a prehľadné rozhranie. V hornej časti je panel otázok, ktoré sa menia podľa odpovedí používateľa. Ten zadáva svoje vstupy pomocou dvoch tlačidiel, Áno alebo Nie. Ak sa používateľ dopracuje ku konkrétnemu riešeniu danej situácie, tak mu aplikácia automaticky zobrazí zoznam možných riešení vo forme príslušných príspevkov. Posledná sekcia je venovaná zoznamu telefónnych čísel na linky záchranných zložiek vo viacerých krajinách sveta. Zoznam je abecedne zoradený podľa názvov krajín, ktoré sú v anglickom jazyku. Kliknutím na krajinu sa zobrazí zoznam čísel rozdelený do

štyroch hlavných kategórií, vid'. obr.4.



Obr. 4 Zoznam telefónnych čísel na linky záchranných zložiek vo viacerých krajinách sveta.

Jednotlivé čísla reprezentujú tlačidlá, ktoré odkazujú na systémovú funkciu zariadenia umožňujúcu okamžité vytáčanie daného čísla.

### C. Testovanie

Vývoj webovej aj mobilnej aplikácie prebiehal formou špirálového životného cyklu, ktorý prebieha po špirále a každá etapa sa opakuje v jednotlivých cykloch. Výhodou je, že vývoj začína malým jadrom – pilotným projektom, ktorý sa po každom cykle rozširuje. Vždy po ukončení jedného celku boli dosiahnuté výsledky konzultované s externým mentorom. Na to plynule nadväzovalo interné testovanie z pohľadu funkcionality a praktickej použiteľnosti aplikácie. Externé testovanie prebiehalo prostredníctvom úzkej skupiny, ktorú tvorili členovia dobrovoľníckych organizácií alebo lekári. Komunikácia s nimi prebiehala vo forme riadenej diskusie a ich názory boli zaznamenané pomocou dotazníka. Hlavnou pripomienkou z ich strany bola aktuálnosť prezentovaných údajov, kde došlo k určitým zmenám pri niektorých typoch ošetrovania počas procesu vývoja aplikácie. Pozitívny ohlas bol zaznamenaný pre možnosť okamžitého vytáčania čísel záchranných zložiek; mnoho ľudí si v strese nemusia spomenúť na tieto čísla.

## III. ZÁVER

Vytvorená aplikácia ako podporné riešenie v situáciách vyžadujúcich prvú pomoc môže byť chápaná ako prvý prototyp, ktorého zavedenie do plnohodnotnej prevádzky si vyžaduje splnenie ďalších faktorov, ako získanie kvalitných informácií pre bázu znalostí; kontinuálne konzultácie s tímom odborníkov; dôkladné testovanie napr. členmi záchranných zložiek, atď. V prípade intenzívnej kolaborácie s príslušnými organizáciami by bolo možné rozšíriť diagnostický systém pre ďalšie kritické situácie, prípadne upraviť proces fungovania inferenčného mechanizmu do efektívnejšej podoby. Ďalšou výzvou by bol dizajn, aby aplikácia dokázala konkurovať iným riešeniam dostupným prostredníctvom on-line trhoviska Google Play pre Android aplikácie.

Zároveň však vytvorená mobilná aplikácia ukazuje, aký veľký potenciál existuje pre podobný typ riešení na Slovensku. Nasadenie vhodných IKT služieb do zdravotníctva môže priniesť nielen úspory na strane poskytovateľov, ale takisto uľahčenie prístupu k zdravotníckym službám zo strany bežných ľudí.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## REFERENCIE

- [1] B.G. Buchanan, E.H. Shortliffe, "Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project", Reading, MA: Addison-Wesley, 1984.



# Efektívnejšie Využívanie Výrobného Opadu Prostredníctvom Jednoduchej Aplikácie

*František BABIČ, Karol KOČIŠ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, karol.kocis.2@student.tuke.sk

**Abstrakt** — Článok je venovaný analýze výrobného procesu v reálnej firme so zameraním na možnú efektívnejšiu realizáciu prostredníctvom zmeny v podprocese evidencie a skladovania odpadu z výroby. Na tento účel bola navrhnutá a implementovaná jednoduchá aplikácia, ktorej cieľom bolo práve zaviesť pozitívnu zmenu do tohto podprocesu za účelom minimalizovať množstvo odpadu na sklade a zároveň maximalizovať jeho využiteľnosť v ďalšej výrobe. Výsledné riešenie bolo nasadené v reálnej prevádzke firmy; následne vyhodnotenie ukázalo zníženie objemu dlhodobého skladovania odpadových plechov na sklade a úsporu času zo strany zainteresovaných pracovníkov.

**Kľúčové slová** — výrobný proces, optimalizácia, aplikácia

## I. ÚVOD

Analýza procesov vo firme predstavuje dôležitú súčasť ich pochopenia a následnej optimalizácie. Každá firma by mala poznať nielen svoje hlavné procesy, ale aj vedľajšie alebo podporné, aby dokázala identifikovať miesta spôsobujúce neplánované prestoje a omeškania. Jednou z možností, ako priniesť zlepšenia do procesov vo firme, je nasadenie vhodného softvérového riešenia práve na základe výsledkov procesnej analýzy.

Manažéri sa v dnešnej dobe zaoberajú zlepšovaním procesov, ich plynulosťou a výkonnosťou. Na to, aby sa dali procesy zlepšovať, je potrebné ich mapovať a analyzovať. Proces je definovaný ako „tok práce, ktorý postupuje od jedného pracovníka k druhému a v prípade väčších procesov z jedného oddelenia do druhého.“ [1, 2] Z výrobného hľadiska pod pojmom proces rozumieme podľa Šaderovej „súhrn pracovných, technologických a prírodných procesov, ktorých pôsobením je pracovný predmet postupne premieňaný vo výrobok.“ [3] Každý proces má svoj určený počet krokov; bod, v ktorom začína a bod, v ktorom končí. Neexistujú všeobecné pravidlá, ako by mal vyzerat' proces v podniku; pre každý podnik sú jeho procesy špecifické v závislosti od predmetu jeho činnosti. Dôležitým faktorom je podchytiť všetky procesy a operácie v procesoch.

Na to, aby firma mohla optimalizovať svoje procesy, musí dobre poznať svoj reálny a základný materiálový tok vo firme. Materiálový tok charakterizuje výrobu v priestore a čase. „Je to riadený pohyb materiálu, surovín, polovýrobovkov, rozpracovaných výrobkov a médií.“ [4] Ideálna optimalizácia materiálového toku pozostáva z nasadenia riešení, ktoré zabezpečia nepretržitý materiálový tok s využitím minimálnych nákladov, spotreby času a ľudských zdrojov.

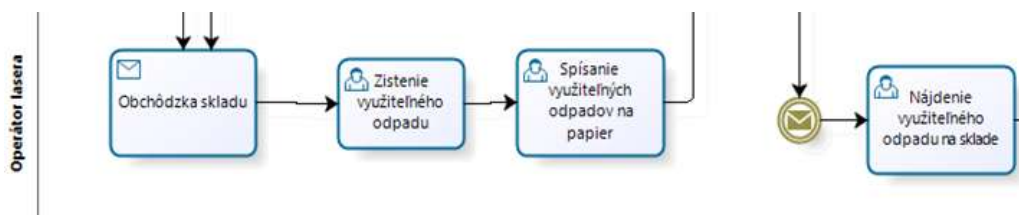
## II. OPTIMALIZÁCIA VÝROBNÉHO PROCESU

### A. Analýza Súčasného Stavú

Výrobný profil analyzovanej firmy je založený na požiadavkách zákazníkov; jednu z hlavných výrobných činností predstavuje pálenie plechov pomocou lasera. Na tomto pracovisku pôsobia dva typy pracovníkov: manažér pracoviska, ktorý má na starosti tvorenie programov na pálenie, objednávanie materiálu, prijímanie objednávok a ich následnú realizáciu; a operátori, ktorí pomáhajú pri príprave plechu na pálenie, triedení vypálených kusov a príprave

na expedíciu. Objednanie materiálu sa realizuje vtedy, ak príde objednávka zo strany zákazníka a na sklade chýbajú požadované plechy. Pri striedaní zmien operátori lasera si musia povedať, ktoré objednávky treba poobede dokončiť, kde sa nachádzajú jednotlivé plechy a využiteľné odpady. Čas venovaný tejto výmene informácií by bolo možné využiť na pálenie ďalších kusov plechu.

Časť vytvoreného procesného modelu pre toto laserové pracovisko je na Obr.1.



Obr. 1 Časť procesného modelu popisujúca práve činnosti, ktoré vykonáva operátor na začiatku celého procesu pálenia.

### B. Návrh Zmeny

Pomocou vytvoreného procesného modelu boli identifikované oblasti, v rámci ktorých by zavedenie vhodných zmien umožnilo optimalizovať fungovanie celého procesu. Jednou z týchto oblastí bola práve evidencia a skladovanie využiteľného odpadu z výroby. Firma nemala zavedenú žiadnu evidenciu tohto odpadu a celý proces zisťovania a využitia vo výrobe bol zbytočne komplikovaný a časovo náročný. Operátor musel vždy najprv osobne zistiť, aké odpady má na sklade a následne, ak ich bolo možné využiť vo výrobe, tak ich znova na sklade nájsť. Takisto manažér pracoviska nemal k dispozícii aktuálny stav uložených využiteľných odpadov; bol odkázaný na informácie od operátora. Všetky tieto činnosti samozrejme vyžadovali čas, ktorý by mohol byť investovaný priamo do výrobného procesu.

Návrhom ako zaviesť pozitívnu zmenu do oblasti využiteľnosti a skladovania výrobného odpadu bolo vytvorenie jednoduchej aplikácie, ktorá by slúžila nielen ako evidencia tohto odpadu, ale zároveň by ponúkala aktuálne informácie všetkým zainteresovaným pracovníkom. Manažérovi by pri tvorbe programov poskytla informácie, aké využiteľné odpady má k dispozícii, a ktoré môže zahrnúť do programu pálenia. Operátor nebude musieť neustále chodiť do skladu a manuálne kontrolovať aký je stav, a kde je ktorý odpadový plech uložený. Priamo po vypálení plechu a vytvorení odpadu by zapísal do aplikácie jeho vlastnosti, rozmery, umiestnenie a odložil by ho do skladu. V prípade, ak by daný plech bolo možné použiť v ďalšom pálení, tak by si pomocou aplikácie vyhľadal jeho umiestnenie a doniesol ho zo skladu. Pri zisťovaní stavu využiteľných odpadov by už manažér nemusel chodiť za operátorom, aby mu overil aktuálny stav, ale pohodlne z pracovného miesta by sa prihlásil do aplikácie a overil si, ktoré využiteľné odpady sú práve na sklade.

Takýmto prístupom by sa skrátil čas pri zisťovaní aktuálneho stavu využiteľných odpadov a zefektívnila by sa príprava na pálenie. Doterajší proces firmy by táto zmena nenarušila. Pri striedaní zmien by si operátori vymenili informácie o výrobnom pláne a objednávkach, ktoré sa musia dokončiť. Nasledujúca zmena by si v aplikácii vyhľadala jednotlivé plechy a využiteľné odpady. Tým pádom by už nemuseli investovať čas do manuálnej obhliadky skladu s odpadom.

### C. Návrh Aplikácie

Na základe výsledkov procesnej analýzy, identifikovaných požiadaviek zo strany pracovníkov firmy a analýze dostupných softvérových riešení bol spracovaný návrh vlastnej aplikácie s prepojením na lokálnu databázu. Zo strany firmy boli stanovené tieto požiadavky:

- pridanie a výber plechu z evidencie,
- prehľad o umiestnení využiteľných odpadov,
- rýchle získanie a predávanie informácií,
- kontrola hospodárenia s využiteľnými odpadmi,
- zefektívnenie práce,
- predbežný aktuálny stav využiteľných odpadov na sklade,
- jednoduché ovládanie aplikácie.
- nasadenie aplikácie na existujúcom PC.

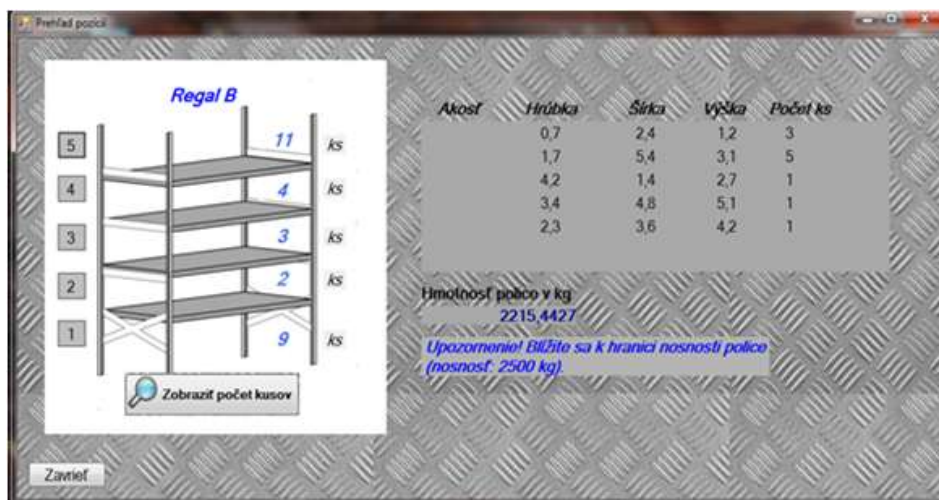
Výhodou vlastnej aplikácie je práve fakt, že bola navrhnutá na mieru, aby splnila špecifické požiadavky firmy. Ponúka základné funkcie na evidenciu využiteľného odpadu, prehľad umiestnenia a voľných pozícií v regáli a zobrazenie aktuálnej ceny, hmotnosti a počtu kusov

využitelných odpadov. Keďže ide o výrobnú firmu, operátori lasera nepoužívajú komplexnejšie programy, ktoré by si vyžadovali IT skúsenosti a znalosti. Plánovaná aplikácia preto nebude náročná na používanie; ponúkne jednoduché úkony s pridaním a vyberaním plechov z databázy. Z tohto dôvodu zaučenie pracovníkov trvalo len niekoľko minút.

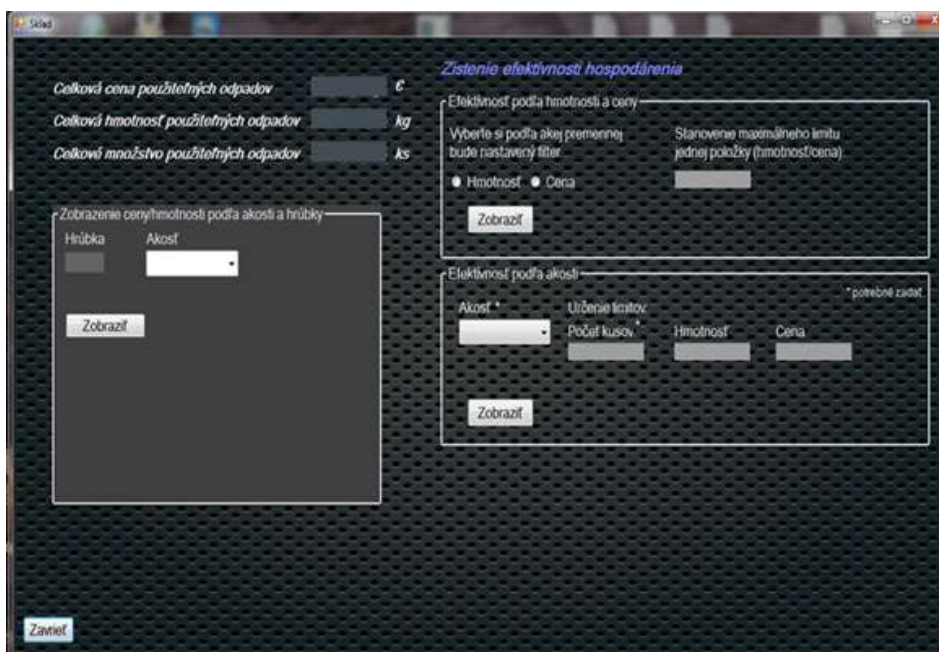
#### D. Výsledok a Testovanie

Implementácia aplikácie prebiehala v niekoľkých fázach; pri prezentácii čiastkových verzií vo firme boli vznesené návrhy na zlepšenia, ako napr. funkcie zistenia efektívnosti hospodárenia. Cieľom tejto funkcie bolo, aby sa predišlo nechcenému stavu využitelných odpadov na sklade napr. v prípade zhromažďovania sa využitelných odpadov jednej akosti alebo hrúbky. Takisto pred reálnym nasadením boli realizované úpravy typu: zmena meracej jednotky z metrov na milimetre, možnosť pridávania ďalších akostí do zoznamu akostí, prednastavené hodnoty niektorých textových polí.

Výsledná verzia umožňuje používateľovi vyhľadať požadovaný plech s možnosťou okamžitého výberu pri použití daného odpadu; pridať nový využitelný plechový odpad do databázy; zobrazenie plechov podľa akosti a prehľad voľných pozícií na regáli; aktuálny stav skladu (hmotnosť, cena, počet kusov), zistenie efektívnosti podľa rôznych charakteristík a zistenie ceny a hmotnosti podľa hrúbky a akosti; cenník: úprava cien plechov. Ukážka niektorých okien aplikácie je na nasledujúcich obrázkoch.



Obr. 2 Okno obsahujúce informácie o umiestnení odpadu na sklade.



Obr. 3 Okno obsahujúce informácie o aktuálnom stave odpadu na sklade.

Pred začatím testovania aplikácie prebehla vo firme inventúra súčasného stavu využiteľných odpadov na sklade. Jednotlivé regály boli označené písmenami tak, ako je to uvedené v aplikácii. Samotné testovanie prebiehalo v pracovnej dobe počas normálneho fungovania firmy. Operátori spisovali využiteľné odpady priamo po vypálení tabule. Manažér využíval aplikáciu vtedy, keď vytváral program na pálenie plechov. Prihlásil sa do aplikácie a zistil, či sa v sklade nachádza požadovaný odpad. Takto už nemusel rušiť operátora a operátor bol oslobodený od obchôdzky skladu zakaždým, keď manažérovi chýbal aktuálny stav využiteľných odpadov. Spokojnosť pracovníkov s aplikáciou bola zisťovaná pomocou dotazníka. Okrem toho boli vyhodnotené dva dôležité ukazovatele a to objem využiteľných odpadov na sklade a čas potrebný na získanie aktuálneho stavu uložených odpadov. V prvom prípade došlo počas testovania aplikácie k zníženiu objemu uložených odpadov o 10% za jeden mesiac. Aplikácia síce neovplyvnila priamo technologický čas pálenia, ale zefektívnila výmenu informácií medzi operátormi a pripravenosť na pálenie. Počas testovania boli vytvárané ďalšie odpady, ale sa aj rýchlejšie spotrebovali, lebo operátori mali väčší prehľad, kde sa čo nachádza a aký odpad majú k dispozícii pre danú akosť plechu. Taktiež mohli predísť prebytočnému kumulovaniu odpadu jednej akosti alebo hrúbky. V druhom prípade došlo k podstatnej redukcii času, keďže manažér už nemusel čakať, pokiaľ mu operátor zistil, aké využiteľné odpady sú na sklade, ale pomocou aplikácie si jednoducho vyhľadal potrebné informácie, a hneď mohol zahrnúť využiteľný odpad do programu. U operátorov sa tento čas skrátil o obchôdzky po sklade. Samozrejme, určitý časový úsek musel byť aj naďalej venovaný odmeraniu a zapisaniu odpadu do aplikácie, ale v porovnaní s predchádzajúcim stavom išlo o zanedbateľné hodnoty.

### III. ZÁVER

Na základe výsledkov testovania je možné konštatovať, že vytvorená aplikácia splnila svoj účel a umožnila efektívnejšiu realizáciu procesu evidencie a následnej využiteľnosti odpadu vo výrobe. Pred zavedením aplikácie sa objavili pripomienky zo strany príslušných pracovníkov, že ďalší „krok“ len skomplikuje ich pracovné činnosti. Počas fázy testovania sa už tieto negatívne postoje stratili, pretože aj samotní pracovníci sa presvedčili o vzniknutých pozitívach: úbytok niektorých úkonov, skrátenie času pri zisťovaní stavu využiteľného odpadu a efektívnejšie využitie výrobného odpadu. Proces evidencie výrobného odpadu je možné aj naďalej vylepšovať, napr. vo forme zavedenia čiarových kódov na regály a plechy, ktoré by uľahčilo prácu manuálneho zapisovania údajov do aplikácie. Vytlačené čiarové kódy by sa nalepovali na plechy a pomocou terminálu by sa zapisovali do databázy. Niektoré využiteľné odpady nemajú obdĺžnikový tvar, preto ďalšou pridanou hodnotou by mohlo byť to, že do aplikácie by sa zapisoval tvar týchto využiteľných odpadov.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] H. Harausová, „Procesné prístupy v manažerstve kvality“, Prešov: Prešovská univerzita, 2012.
- [2] J. Daněk, J. Pavliška, „Technológie ložných a skladových operácií“, II. vydanie, Ostrava: VŠB – Technická univerzita Ostrava, 2002.
- [3] J. Šaderová, J. Boroška, „Výrobné procesy“, 1. vydanie, Košice: TU – BERG, edičné stredisko/AMS, 2005.
- [4] P. Šebej, „Optimalizácia, topológia tokov vo výrobných procesoch“, in: *Posterius*, 5(8), 2012, str. 1-5.

# Klasifikácia Emailov Pomocou Vhodných Metód Dolovania v Dátach

František BABIČ, Petra ŽULEVIČOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, petra.zulevicova@student.tuke.sk

**Abstrakt** — Komunikácia prostredníctvom emailov patrí medzi najstaršie on-line spôsoby komunikácie, avšak nárast intenzity používania priniesol aj rozšírenie nevyžiadanej pošty (SPAMu). Tento fenomén modernej komunikácie naštartoval vývoj a aplikáciu rôznych algoritmov a metód, cieľom ktorých bolo identifikovať kľúčové prvky charakteristické pre nevyžiadanú poštu a následne ich využiť na tvorbu nástrojov ako napr. SPAM filtre. Naším cieľom bolo prispieť k riešeniu tejto problematiky a prostredníctvom dvoch voľne dostupných dátových množín definovať model SPAMu, pozostávajúci práve z extrahovaných kľúčových charakteristík.

**Kľúčové slová** — email, SPAM, dolovania v dátach

## I. ÚVOD

Jedna z najčastejších činností vykonávaná pri práci s počítačom je písanie resp. čítanie emailov. Každý používateľ emailovej komunikácie denne prečíta a odošle niekoľko emailov, z ktorých pomerne veľké množstvo predstavuje práve SPAM. Stručne povedané, SPAM je taký email, ktorý bol odoslaný príjemcovi bez jeho predchádzajúceho súhlasu, vo väčšine prípadov s pochybným obsahom a totožnosť odosielateľa je anonymná [4]. S narastajúcim počtom SPAM emailov sa poskytovatelia emailových služieb rozhodli zamerať na zníženie a odstránenie neustále sa zvyšujúceho počtu takýchto správ. Analýzou obsahu emailovej správy je možné pomerne efektívne určiť, či sa jedná o legitímnu poštu (HAM) alebo poštu s klamlivým a zavádzajúcim obsahom (SPAM). Nástroje na filtrovanie nevyžiadanej pošty fungujú na báze naprogramovaných pravidiel, ktoré obsahujú kľúčové slová alebo charakteristiky získané z hlavičky a tela emailu. SPAM zo sebou prináša riziko, pretože okrem neškodného reklamného obsahu môže obsahovať aj nebezpečné odkazy na škodlivé stránky alebo prílohy obsahujúce škodlivý softvér.

Analytický proces bol realizovaný pomocou softvérového nástroja RapidMiner v súlade s metodológiou CRISP-DM, ktorá predstavuje všeobecný rámec na riešenie úloh dolovania v dátach [1, 2].

## II. PROCES ANALÝZY EMAILOV

### A. Cieľ Analýzy

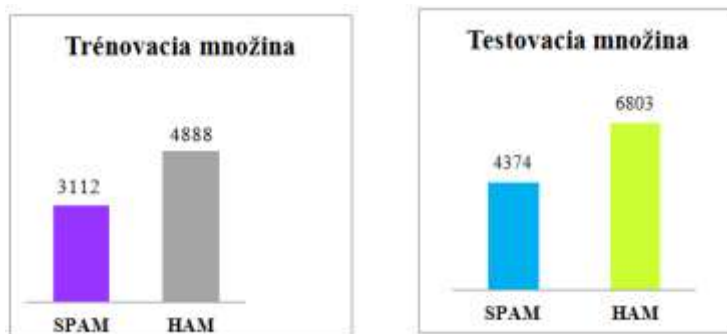
Elektronická pošta stala dôležitým marketingovým nástrojom pre budovanie a udržiavanie vzťahov so zákazníkmi. E-mailový marketing sa stal obľúbenou formou na propagáciu výrobkov a služieb, rozvoj celosvetovej siete Internet umožnil oslovenie širokej cieľovej skupiny zákazníkov. Nárast využívania e-mailového marketingu vyvolal rast nevyžiadaných emailov známych ako SPAM. Tvorcovia nevyžiadaných správ vyvíjajú čoraz viac sofistikovanejšie spôsoby na prekonanie tradičných SPAM filtrov. Prešli od jednoduchých textových správ k obrázkovému SPAMu; SPAMu s nevyžiadanou prílohou, čo predstavuje určitú záťaž pre elektronickú komunikáciu a taktiež si to vyžaduje vyšší výpočtový výkon na odlišenie nevyžiadanej pošty od legitímnych správ. E-mailový marketing nie je jedinou formou SPAMu, ktorý sa dostane do našich elektronických schránok. Za SPAM možno považovať najmä také emaily, ktorých obsah sa týka pornografie, predaja on-line liekov, výhry v lotérii a iné. Najnovšou taktikou pre maskovanie obsahu SPAMu v emailoch je nahrádzanie písmen symbolmi, ktoré sú pre človeka stále ľahko čitateľné a zároveň nie sú SPAM filtermi označené

ako nevyžiadaná pošta.

Naším cieľom bolo vytvoriť klasifikačné modely na základe dvoch vstupných dátových množín. Pomocou viacerých experimentov sme overili schopnosť vytvorených modelov čo najpresnejšie klasifikovať emaily ako SPAM alebo HAM. Následne, model s najlepšou predikčnou schopnosťou bol použitý na extrakciu kľúčových charakteristík s cieľom vytvoriť všeobecný model SPAMu.

### B. Popis Dátových Množín

Prvá dátová množina obsahovala viac ako 19 tisíc emailov, z ktorých 8 tisíc predstavovalo trénovaciu vzorku. Cieľový atribút *target* nadobúdala hodnoty typu *Yes* alebo *No*. Hodnota *Yes* určuje, že daný email je SPAM a hodnota *No* hovorí, že ide o legitímnu správu. Distribúcia hodnôt cieľového atribútu v trénovacej a testovacej množine je zobrazená na Obr.1.



Obr. 1 Distribúcia hodnôt cieľového atribútu v dátovej sade č.1.

Každý email v dátovej množine bol popísaný prostredníctvom 832 ďalších atribútov, ktoré predstavujú zozbierané charakteristiky (frázy) vyskytujúce sa či už v SPAM alebo HAM emailoch. Tieto binárne atribúty určili, či sa daná charakteristika v emailoch nachádza alebo nie. Túto pomerne veľkú množinu atribútov bolo možné rozdeliť na dve hlavné skupiny podľa toho, v ktorej časti emailu sa nachádzali:

Hlavička emailu, napr.

- MISSING\_OUTLOOK\_NAME - email vyzerá ako poslaný z MS Outlook, ale nie je.
- DOMAIN\_4U2 - názov domény obsahuje „4U“-„pre teba“.
- ADVERT\_CODE2 - pole: Predmet obsahuje slogan/značku inzerátu.
- SUBJ\_HAS\_SPACES - pole: Predmet obsahuje veľa medzier.
- DATE\_IN\_FUTURE\_12\_24 - pole: Dátum je 12 až 24 hodín po dátume prijatia.
- FROM\_HAS\_MIXED\_NUMS - pole: Odosielateľ obsahuje mix písmen a čísiel.

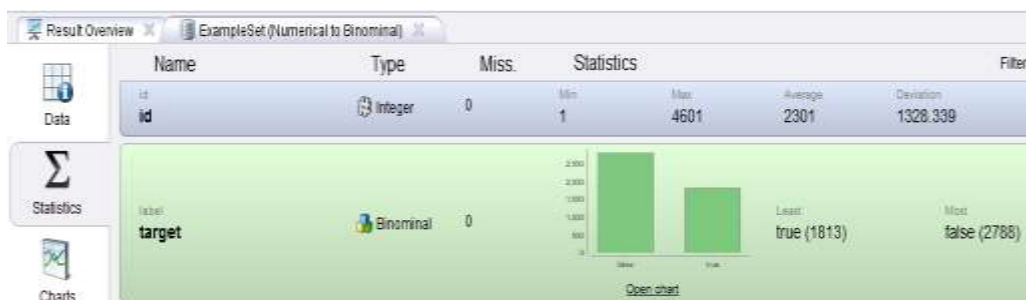
Telo emailu, napr.

- CTYPE\_JUST\_HTML - čistý HTML email, bez textovej verzie.
- LINES\_OF\_YELLING - obsahuje celý riadok veľkými písmenami.
- UPPERCASE\_50\_75 - telo emailu má 50 až 75% veľkých písmen.
- ALL\_CAP\_PORN - slovo PORNO s veľkými písmenami.
- CLICK\_BELOW - žiada kliknúť » sem.
- EXCUSE\_10 - ospravedlnenie „ak si viac neželáte prijímať ďalšie...“.
- LOW\_PRICE - najnižšia cena.
- SATISFACTION - spokojnosť zaručená.
- US\_DOLLARS\_3 - nigérijský SPAM výraz.
- HTML\_FONT\_COLOR\_BLUE - kód farby písma v HTML je modrý.
- WEIRD\_PORT - používa neštandardný port pre http.

Popis a pochopenie atribútov sme čerpali z relevantnej literatúry, napr. [3].

Dátová množina č.2 predstavovala menšiu vzorku, konkrétne 4 601 emailov. Distribúcia hodnôt cieľového atribútu v tejto množine je zobrazená na Obr.2. Táto dátová množina obsahovala 48 atribútov typu *word\_freq\_SLOVO* zodpovedajúce % výskytu daného slova v emailovej správe. Okrem toho sa v nej nachádzalo aj 6 atribútov typu *char\_freq\_ZNAK*, ktoré vyjadrujú frekvenciu výskytu znaku v správe a 3 atribúty typu: priemerná dĺžka neprerušenej sekvencie veľkými písmenami, dĺžka najdlhšieho neprerušeneho sledu veľkých písmen a celkový počet veľkých písmen v emailovej správe.





Obr. 2. Distribúcia hodnôt cieľového atribútu v dátovej sade č.2.

### C. Experimenty

Dátová množina č.1 neobsahovala žiadne chybné alebo chýbajúce hodnoty, preto sme sa v tejto fáze zamerali skôr na atribúty a ich vplyv na ten cieľový, čiže úspešnosť klasifikácie. Pomocou histogramov a základných štatistík sme sa rozhodli z danej množiny odstrániť tie atribúty, ktoré sa v emailoch nenachádzali ani raz, alebo iba raz, maximálne dvakrát. Pôvodný počet atribútov 832 sa znížil o 373. Napriek tejto operácii, príslušné klasifikačné experimenty nepriniesli očakávané výsledky. Preto sme sa rozhodli ďalej pracovať s množinou atribútov, konkrétne sme využili hodnotiacu funkciu Ginni-index na určenie atribútov s najvyššou váhou k cieľovému. V tomto prípade bolo vždy nutné zvoliť hodnotu koeficientu  $k$ , ktorý určoval počet vybraných atribútov. Experimentálne sme overili viaceré kombinácie koeficientu  $k$  a aplikovanej metódy dolovania v dátach s cieľom dosiahnuť čo najpresnejší klasifikačný model. Pre rozhodovací strom bola najvhodnejšia hodnota  $k$  rovná 190, v prípade algoritmu  $k$ -nn 155 a pre neurónové siete 200. Tieto hodnoty určujú, koľko vstupných atribútov bolo v danom prípade použitých na vytvorenie klasifikačných modelov. V druhom prípade bola k dispozícii podobne kvalitná dátová množina, ktorá neobsahovala žiadne chybné alebo chýbajúce hodnoty. Dôležitým krokom bolo určiť práve vhodné rozdelenie na tréningovú a testovaciu množinu, kde rôzne alternatívy boli postupne experimentálne overené a vyhodnotené.

Vo fáze modelovania boli aplikované na obe dátové množiny tri metódy dolovania: rozhodovacie stromy, algoritmus  $k$ -najbližších susedov a neurónové siete. Keďže cieľom bolo nielen čo najpresnejšie klasifikovať emaily do dvoch cieľových kategórií, ale aj extrakcia kľúčových charakteristík pre SPAM emaily, v konečnom dôsledku boli rozhodovacie stromy práve tým najvhodnejším prístupom pre ich jednoduchú reprezentáciu rozhodovacích pravidiel.

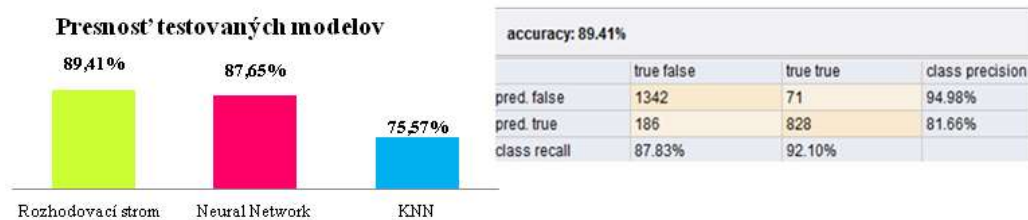
Vytvorené modely boli vyhodnotenie nielen na základe % úspešnosti, ale aj schopnosti správne klasifikovať emaily do dvoch cieľových tried. Na prvej dátovej množine boli dosiahnuté najlepšie výsledky pomocou rozhodovacích stromov, čo dokumentuje aj Obr.3.



Obr. 3 Presnosť najlepších modelov na testovacej vzorke, dátová množina č.1.

Testovaný model správne klasifikoval ako SPAM 4 252 emailov a ako legitímnu poštu 6 720 emailov. Model nesprávne klasifikoval 205 emailov, čo predstavuje celkovú chybu klasifikácie vo výške 1,83 %.

V prípade druhej dátovej množiny bolo uskutočnených viacero experimentov, najvyššie hodnoty presnosti priniesli modely pri rozdelení približne 50% vzorka na tréningovanie a 50% vzorka dát na testovanie. Výber emailov do oboch vzoriek prebiehal náhodne, čo v niektorých prípadoch mohlo ovplyvniť dosiahnuté výsledky. Presnosť klasifikácie najlepších modelov u všetkých troch metód je zobrazená na Obr.4 spolu s popisom správnosti klasifikácie modelu založenom na rozhodovacích stromoch.

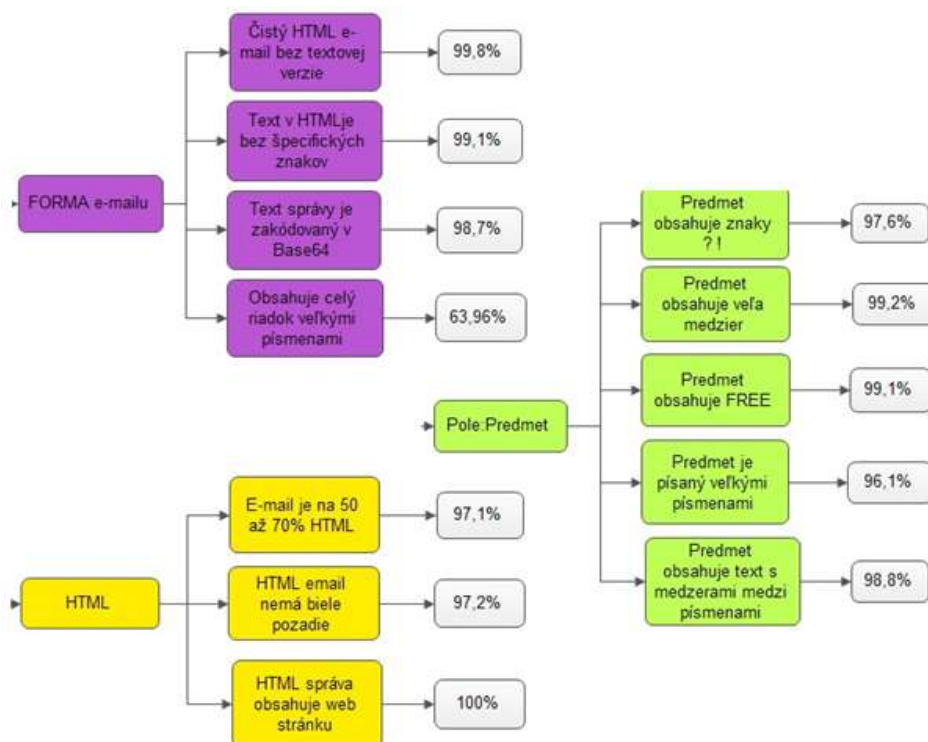


Obr. 4 Presnosť najlepších modelov na testovacej vzorke, dátová množina č.2.

V testovacej množine sa nachádzalo 899 SPAM emailov a 1 528 legitímnych správ, testovaný model rozhodovacieho stromu správne klasifikoval ako legitímnu poštu 1342 emailov a ako SPAM 828 emailov. Chyba klasifikácie bola 10,59%, čo predstavuje 257 zle klasifikovaných emailov. V tomto prípade boli dosiahnuté výsledky o niečo nižšie, čo bolo mohlo byť spôsobené iným charakterom atribútov popisujúcich jednotlivé emaily alebo nižším počtom atribútov reprezentujúcich kľúčové slová. Takisto veľký vplyv malo rozdelenie dátovej množiny na tréningovú a testovaciu vzorku, kde pri niektorých experimentoch došlo k situácii, že niekoľko emailov nebol vytvorený model schopný zaradiť do žiadnej kategórie. Táto situácia mohla byť spôsobená tým, že daný typ emailu s rôznou frekvenciou výskytu slov a znakov sa v tréningovej množine nevyskytoval, tým pádom model nebol naučený, či daný email má klasifikovať ako SPAM alebo ako legitímnu poštu.

#### D. Všeobecný Model SPAMu

Dosiahnuté výsledky nám poskytli dostatok informácie o tom, ktoré charakteristiky, slová a znaky sú špecifické pre SPAM emaily. Všeobecný model SPAMu pozostáva zo sady najvýznamnejších SPAM charakteristík vybraných z oboch dátových množín na základe rozhodovacích pravidiel získaných z výsledných testovacích modelov rozhodovacieho stromu. Vybrané charakteristiky boli kategorizované do tried podľa miesta výskytu v hlavičke alebo tele emailu. Aby správa bola vyhodnotená ako SPAM, musí email obsahovať viacero SPAM charakteristík, na základe ktorých sa správe prideli SPAM skóre. Ukážka niektorých častí tohto modelu je na Obr.5.



Obr. 5 Ukážka všeobecného SPAM modelu na základe realizovaných experimentov.



### III. ZÁVER

Objem emailov, ktoré denne dostávajú do emailových schránok stále rastie, čím rastie aj počet SPAM emailov. Z tohto dôvodu rastie aj význam detekcie, ako pre firemné tak aj pre verejné inštitúcie. Väčšina SPAM filtrov poskytuje svoje služby na základe naprogramovaných pravidiel, ktoré triedia prichádzajúcu poštu do priečinkov, pričom SPAM emaily sú odstránené na základe kľúčových slov resp. charakteristík emailu získaných z hlavičky a tela emailu. Charakteristiky SPAM emailov sa v priebehu času menia, preto je dôležité, aby pravidlá pomocou ktorých sa triedi SPAM od legitímnych správ boli aktualizované; inak bude softvér pre detekciu SPAMU so zastaranými pravidlami náchylný k chybám, čo bude viesť k prepusteniu SPAMu do emailového adresára.

Výsledky experimentov, ktoré sme realizovali na oboch dátových množinách ukazujú, že z použitých metód sú práve rozhodovacie stromy tým najúspešnejším klasifikátorom. Na základe vytvorených klasifikačných modelov sme zostavili všeobecný model SPAMu, ktorý môže slúžiť ako predloha pre špecifikáciu antiSPAMových pravidiel.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] P. Pavel, „Data Mining“, diel I., Univerzita Pardubice, 2006, s.8-18.
- [2] C. Shearer, „The CRIPS-DM Model: The New Blueprint for Data Mining“, in *Journal of Data Warehousing*, 5(4), 2000, s.13-22.
- [3] M. Spykerman, „Typical spam characteristics“, Red Earth Software.
- [4] Yu Szde, „Email spam and the CAN-SPAM Act: A qualitative analysis“, in *International Journal of Cyber Criminology*, 5 (1), 2011, s.715-732.

# On-line Aplikácia pre Komplexný Manažment Taxislužby

*František BABIČ, Lukáš STAROŇ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, lukas.staron@student.tuke.sk

**Abstrakt** — Článok popisuje proces návrhu a implementácie on-line aplikáciu pre komplexný manažment taxislužieb. Požiadavky na aplikáciu boli získané analýzou existujúcich podmienok, dostupných konkurenčných riešení a prostredníctvom spolupráce s konkrétnym subjektom. Aplikácia poskytuje rozhranie pre úplný manažment taxislužby - od evidencie zamestnancov cez zoznamy objednávok až po zobrazenie trás na mape. Takisto podporuje modulárny prístup, čo znamená že základnú množinu funkcií je možné kedykoľvek doplniť pomocou inštalácie vhodných zásuvných modulov. Výsledné riešenie bolo testované v reálnych podmienkach, je schopné základnej prevádzky, ale pre plnohodnotné nasadenie je nutná jeho kustomizácia na konkrétny subjekt a najmä používatelia musia byť patrične zaškolení.

**Kľúčové slová** — aplikácia, taxislužba, modul

## I. ÚVOD

Táto práca bola venovaná návrhu a implementácii on-line aplikácie, ktorá má slúžiť na podporu prevádzky a manažmentu taxislužieb. "Taxislužba je prevádzkovaním osobnej dopravy vozidlami taxislužby ako prepravy jednotlivých cestujúcich alebo skupiny cestujúcich do cieľového miesta podľa zmluvy o preprave osôb. Dopravca má prevádzkovú povinnosť v rozsahu koncesie, prepravnú povinnosť v rozsahu prepravného poriadku a tarifnú povinnosť podľa tarify." [1]

Cieľom bolo vytvoriť také softvérové riešenie, ktoré v podstatnej miere zjednoduší vonkajšie aj vnútorné procesy taxislužby, jej zamestnancov a zákazníkov. Práca bola motivovaná analýzou súčasnej situácie v oblasti prepravných služieb na území mesta Košice. Vo všeobecnosti bolo možné pozorovať viacero faktorov, z ktorých vyplynula absencia akéhokoľvek rozsiahlejšieho informačného systému, ktorý by riešil túto problematiku, príslušné procesy, atď. Napr. obyvateľom chýbala aplikácia, ktorá by im umožnila ľahko a rýchlo komunikovať nielen s dispečermi ale aj samotnými vodičmi. Na jednom mieste by mali k dispozícii informácie, kde sa daný vodič nachádza, kedy sa dostaví na ich adresu alebo koľko budú platiť. Vodiči a dispečeri zároveň potrebovali rozhranie, pomocou ktorého by si mohli kontrolovať svoje objednávky a ich trasy. Majitelia spoločnosti zas oceňujú, ak v ňom budú schopní manažovať svojich zamestnancov a delegovať na nich prácu. Existencia týchto faktorov podnietila návrh a implementáciu komplexnej aplikácie dostupnej na všetkých typoch inteligentných zariadení.

Je samozrejmé, že na trhoch už existujú riešenia, ktorých podstata je venovaná práve manažmentu taxislužieb, ich vodičom a zákazníkom. Všetky sa v zásade od seba odlišujú určitými parametrami, ako napr. dostupnosťou len v danej krajine. Zákazníci spoločností ako Get Taxi [2], Taxi Magic [3], myTaxi [4] alebo Go-Taxi [5] majú k dispozícii detailný prehľad vozidiel a ich vodičov. Dokážu si prezeráť ich hodnotenia, profily a profilové fotografie, vyťaženosť a iné doplnujúce údaje. Môžu vytvárať objednávky, pri ktorých vidia presnú polohu blížiacieho sa taxíka prostredníctvom GPS. Za prepravu môžu zaplatiť v hotovosti, cez telefón alebo platobnou kartou. Spoločnosti disponujú mobilnými aplikáciami pre smartfóny, ktoré sú k dispozícii prostredníctvom App Store, Google play, Windows Phone a BlackBerry World. Okrem toho je možné všetky činnosti vykonávať aj prostredníctvom webovej stránky s integrovanými mapami. Na druhej strane zase stoja aplikácie, ktoré umožňujú bežným ľuďom zdieľať buď ich alebo dostupné vozidlá, ako napr. Uber [6], čím vytvárajú nechcenú

konkurenciu pre tradične fungujúce taxislužby.

## II. NÁVRH A IMPLEMENTÁCIA APLIKÁCIE

Proces návrhu riešenia bol od počiatkových fáz konzultovaný s minimálne jednou reálnou spoločnosťou prevádzkujúcou taxislužbu. Jednotlivé funkcionality a ich podoba vychádzajú z identifikovaných potrieb budúcich používateľov.

Aplikácia obsahuje zásuvný modul *Správa používateľov* pre kompletný manažment používateľov. Je možné ich pridávať, upravovať a vymazávať. Informácie zadáva používateľ prostredníctvom profilového formuláru. Ten obsahuje povinné a voliteľné polia. Každý používateľ musí mať pridelenú aj používateľskú rolu, ktorú vyberá administrátor s oprávnením na základe funkcie, ktorú má nový člen v aplikácii plniť. Samotný modul, všetky jeho sekcie a každá funkcia v nich sú vybavené kontrolou používateľskej roly a úrovne zabezpečenia. Iba oprávnení používatelia majú teda možnosť prístupu k jednotlivým funkcionalitám. Kontrola úrovne bola zavedená z jedného prostého dôvodu – zabránenie používateľom nižšej úrovne upravovať účty používateľov na úrovni vyššej.

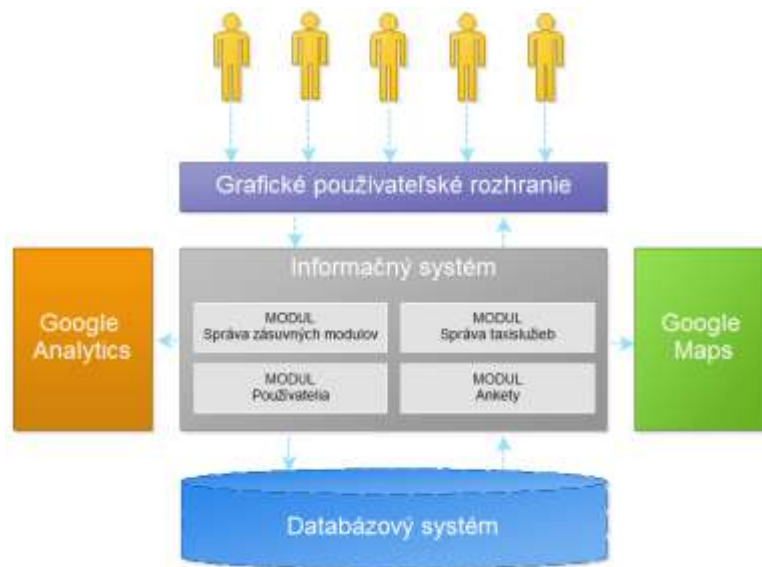
Modul *Administrácia* predstavuje skupinu funkcionalít, na ktorých sú postavené, resp. o ktoré sa opierajú ostatné zásuvné moduly. Administrácia vytvára rozhranie pre prístup iba prihlásených používateľov, ktorí disponujú patričnými používateľskými právami. Modul je zodpovedný za automatické generovanie a dynamické prispôsobovanie sa administratívnej navigácie. Hlavnou úlohou modulu je však najmä vytvorenie priestoru pre zmenu systémových nastavení a možnosť správy všetkých zásuvných modulov. Používateľ s potrebnými právami má prehľad o nainštalovaných moduloch. Vidí, či sú funkčné, alebo je pri nich hlásená porucha. Má možnosť ich inštalovať a odinštalovať.

Modul *Taxislužba* je dôležitou súčasťou celej aplikácie, pretože ponúka možnosť plnej správy taxislužieb, ich zamestnancov, objednávok a rozhrania pre dispečerov. Správca prezerá, pridáva, upravuje a vymazáva taxislužby. Udeľuje im oprávnenie používať systém prostredníctvom prideleného verifikačného kľúča. Každý kľúč sa generuje automaticky a zašle majiteľovi taxislužby, iným ako bežným komunikačným kanálom. Ten jeho vloženie do aplikácie potvrdí svoje kontaktné údaje a sprístupní si celý systém. Takto overený správca taxislužby má možnosť vytvoriť si svoju firemnú štruktúru, ktorej budú zodpovedať jednotlivé role v rámci reálnej taxislužby. Ponúkne sa mu tým možnosť vytvoriť si také zamestnanecké pozície, akými už jeho spoločnosť disponuje, resp. aké si želá. Tieto pozície následne prideluje používateľom, ktorých určuje za svojich zamestnancov. Ďalej sa títo používatelia z pohľadu taxislužby správajú len tak, ako im bolo výberom roly predurčené. Objednávky môžu vytvárať všetci používatelia aplikácie. Dodávajú potrebné povinné údaje (kontaktné údaje, štart, cieľ, počet osôb, batožina, ...) a vyberajú si vodiča, ktorému je zákazka pridelená. Dispečer taxislužby môže navyše prezeráť, upravovať a vymazávať jednotlivé objednávky. Taktiež má možnosť prezeráť všetky aktívne zákazky priamo na mape s trasou, s vyznačenými štartovacími a cieľovými bodmi, všetky farebne odlišené. Zobrazovanie mapy spolu s vytýčenými trasami zákaziek bolo implementované pomocou rozhrania Google Maps, vid'. Obr. 1.



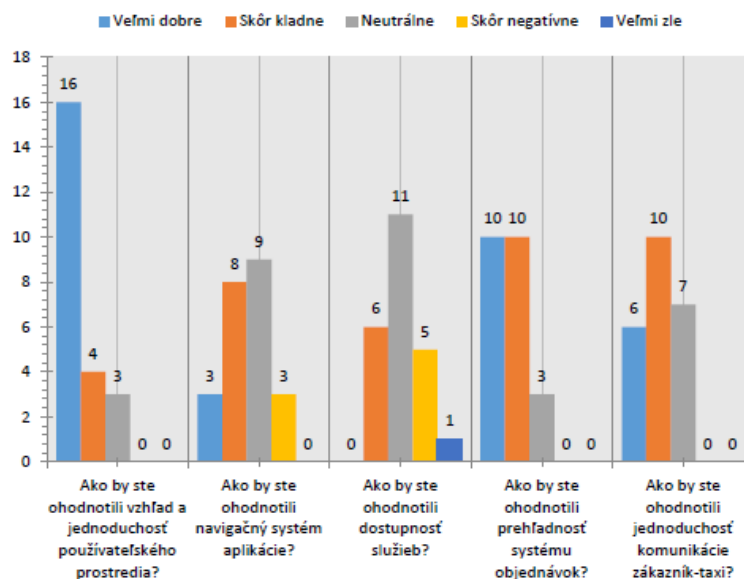
Obr. 1 Ukážka vizualizácie trás objednávok na mape.

Základná vnútorná architektúra aplikácie je zobrazená na Obr.2. Komunikácia s používateľmi prostredníctvom GUI a komunikácia s databázou prebiehajú obojsmerne; používateľ prostredníctvom rozhrania formuluje požiadavky na aplikáciu. Tá ich spracuje, využívajúc funkčnosti svojich integrovaných zásuvných modulov a prezentuje odpoveď rovnakým kanálom ako prišla požiadavka – prostredníctvom GUI. Komunikácia s databázou prebieha obdobným spôsobom; väčšina komunikácie má formu dopytovania v jazyku SQL. Okrem týchto základných spôsobov výmeny informácií prebieha časť komunikácie aj so službami spoločnosti Google. Jedna z nich, tá dôležitejšia, poskytuje rozhranie pre prácu s mapami. Prostredníctvom konkrétneho JavaScript kódu vytvára mapu a jej komponenty, tie následne integruje na presné miesto v aplikácii. Druhá služba od spoločnosti Google má rýdzo informačný charakter. Vytvára komplexný súbor štatistík o návštevníkoch a ich správaní na základe priebežného monitoringu a zberu informácií.



Obr. 2 Ukážka vnútornej architektúry aplikácie.

Aplikácia prešla počas implementácie viacerými fázami interného testovania s cieľom odhaliť nesprávne fungovanie aplikácie v jednotlivých možných prípadoch použitia. Následne bola nasadená do umelej reálnej prevádzky, t.j. na vybranej skupine vozidiel boli testované jednotlivé funkčnosti pre koncových používateľov s cieľom získať spätnú väzbu, návrhy na opravy, prípadne vylepšenia, ktoré by aplikáciu ešte viac priblížili reálnym požiadavkám použitia. Tieto informácie boli získané dotazníkovým prieskumom, výsledky ktorého priniesli celkovú spokojnosť používateľov cez 70%, vid'. Obr.3.



Obr. 3 Ukážka spätnej väzby pomocou dotazníkového prieskumu.

### III. ZÁVER

Výsledkom tejto práce je funkčný prototyp aplikácie, ktorá komplexným spôsobom podporuje interné procesy v taxislužbe. Pomocou modulu taxislužba je možné spravovať zamestnancov, zamestnanecké pozície a všetky evidované objednávky. Okrem toho je k dispozícii integrovaný systém Google Maps, prostredníctvom ktorého je možné monitorovať aktuálne prebiehajúce zákazky pomocou farebne odlišených trás. Zo spätnej väzby od používateľov a z pozorovaní sa vykryštalizovalo niekoľko nedostatkov a príležitostí pre vylepšenia do budúcnosti. Bolo by vhodné doplniť podporu elektronických platieb či vytvoriť prostredie pre účtovné prehľady a štatistiky. Dispečeri by ocenili rozšírenie možností ich rozhrania o nové funkcie ako: zdieľanie vozidiel, automatické odporúčania zmeny stanovíšť a prehľady spoločenských akcií v meste. Rovnako bola identifikovaná potreba zjednotenia prostredia aplikácie, navigácie, komunikačného nástroja, platobného terminálu a pokladnice do jedného uceleného zariadenia.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] Zákon č. 56/2012 Z.z. o cestnej doprave
- [2] Riešenie Get Taxi, <http://www.gettaxi.com/>
- [3] Riešenie Taxi Magic, [https://taximagic.com/en\\_US](https://taximagic.com/en_US)
- [4] Riešenie myTaxi, <https://www.mytaxi.com/home.html>
- [5] Riešenie Go-Taxi, <http://www.go-taxi.biz/>
- [6] Riešenie Uber, <https://www.uber.com>

# Riadenie pre autonómne hydroponické systémy

<sup>1</sup>Rudolf JAKŠA, <sup>2</sup>Jakub SZÁSZ

<sup>1,2</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika  
<sup>1</sup>rudolf.jaksa@tuke.sk, <sup>2</sup>jakub.szasz@student.tuke.sk

**Abstrakt** — V tomto článku je pojednávané o bakalárskej práci ktorá sa zaoberá hydroponickým spôsobom pestovania rastlín a spôsobom udržiavania určitých vlastností prostredia v praxi. Na základe týchto poznatkov je vypracovaný návrh hardwareového vybavenia jedného z typov hydroponických systémov, spolu s návrhom algoritmu pre reguláciu prostredia.

**Kľúčové slová** — hydroponia, fuzzy, kognitívne, mapy, riadenie, pestovanie

## I. ÚVOD

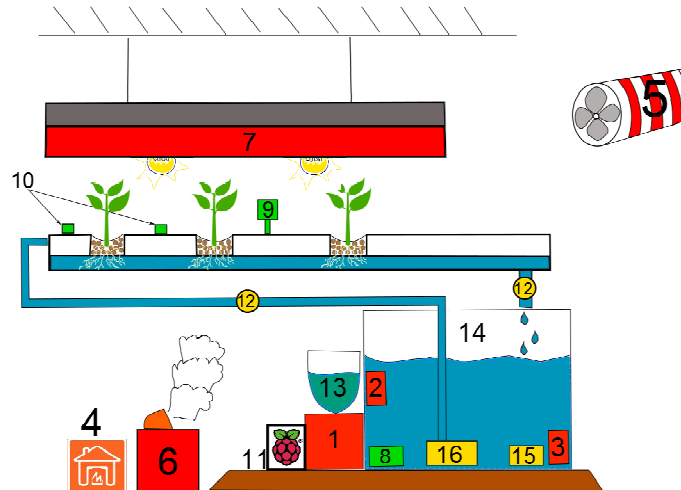
Ľudstvo už od dávna pozná a využíva hydroponický spôsob pestovania plodín na živnom roztoku. Doposiaľ je však veľmi málo rozšírený, či už ako pojem medzi širokou verejnosťou alebo ako pestovateľský postup pre dosiahnutie vyššej kvality i kvantity pestovaného produktu. Dôvodom je to, že si ľudia dokážu vystačiť s množstvom rastlinnej potravy, ktorú si zabezpečia klasickým spôsobom pestovania v zemi. Pri hydroponickom spôsobe pestovania rastlín je potrebné dbať na jeho funkčnosť a neustále kontrolovať stav prostredia. Tieto predpoklady z neho robia výborného kandidáta pre použitie niektorých z metód umelej inteligencie na vytvorenie systému pre jeho autonómnou správu. Cieľom bakalárskej práce bolo mimo iného, navrhnúť schému autonómneho hydroponického systému typu flood & drain spolu so softwarovým vybavením zabezpečujúcim správu prostredia rastlín (aplikácia nazvaná „Life Support“).

## II. HYDROPÓNIA

Ako bolo spomenuté, hydroponia patrí medzi zložitejšie, menej využívané, ale za to výnosnejšie a ekologickejšie spôsoby pestovania rastlín. Pri tomto spôsobe nie je použitá hlina ako základné médium dodávajúce rastlinám potrebné látky, ale je vystriedaná inertným vodozadržujúcim substrátom (napr.: minerálna vlna, kokosové vlákna, keramzit) a živným roztokom plným jednobunkových organizmov, ktorý je ku koreňom rastliny privádzaný, buď sekvenčne alebo bez prerušenia (v závislosti na použitom type systému). Živný roztok by mal disponovať vyváženým a správnym množstvom živín, ktoré sú primiešavané do vody v podobe minerálnych solí. Kvantitatívne znaky živín v roztoku sa prejavujú na stupnici elektrickej vodivosti. V zásade platí, že čím je väčšie množstvo živín v roztoku, tým je vyššia jeho elektrická vodivosť. Pri živnom roztoku dbáme taktiež na hodnotu pH, ktorá sa taktiež ako množstvo živín počas procesu mení. PH živného roztoku musí byť udržiavané v určitom rozsahu, pretože schopnosť rastlín prijímať jednotlivé prvky sa mení v závislosti na ňom. Klasický spôsob pestovania vyžaduje rozhodne menšiu starostlivosť o tieto faktory, pretože zemina má schopnosť samoregulácie niektorých vlastností. Hodnoty pH a elektrickej vodivosti sa nastavujú pomocou pridávania rôznych zmesí na jeho úpravu. Prostredie rastlín ďalej tvorí vzduch, ktorý by mal mať správnu teplotu, relatívnu vlhkosť a obsah CO<sub>2</sub>. Správne hodnoty týchto veličín zabezpečujeme napr.: ohrievačom a ventiláciou, ale aj klimatizáciou, zvlhčovačom, poprípade pridávaním oxidu uhličitého. V neposlednej rade sa všeobecne pri pestovaní kladie dôraz na veľkosť svetelnej intenzity, kvôli procesu fotosyntézy, ktorý je pre rastlinu životne dôležitý. Ako je vidieť, udržiavať hydroponický systém v rovnováhe je proces, kde vystupujú rozličné faktory a rozličné spôsoby pre ich reguláciu. V prípade, že zlyhá kontrola nad takýmto systémom, sú postihnuté rastliny, či už vo forme zníženia produkcie, patogénov, chorôb alebo úmrtia.

### III. IMPLEMENTÁCIA

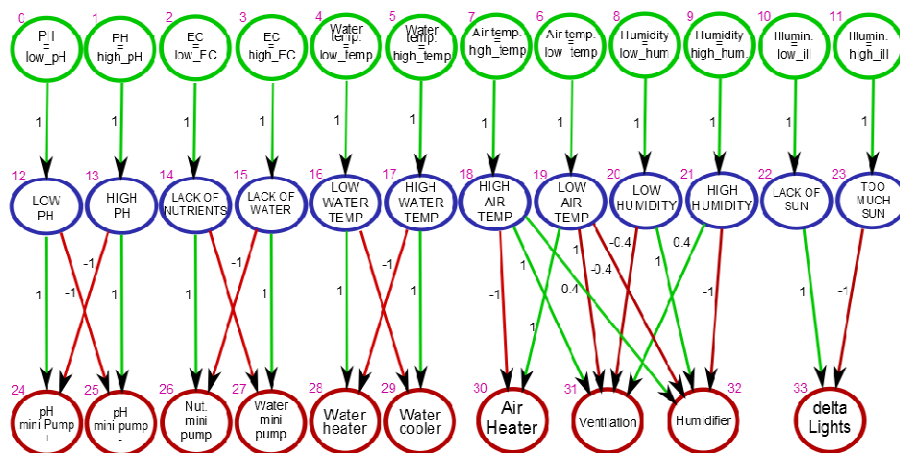
Jedným z cieľov bakalárskej práce bolo vytvoriť model autonómneho hydroponického systému typu flood & drain. Tento typ systému sa vyznačuje tým že živný roztok je ku koreňom rastlín privádzaný pravidelne pričom sa vždy zaplaví a odploví celý koreňový systém. Ďalej uvažujeme, že sa jedná o tzv. indoor-ový štýl pestovania (skleník, miestnosť). Schéma navrhnutého systému je zobrazená na obr.1.



**Obr. 1 – Schéma autonómneho flood & drain systému.**

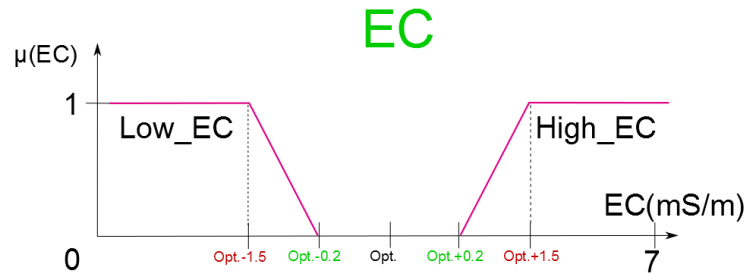
- 1) Mini pumpy, 2) Ohrievač vody 3) Ochladzovač vody, 4) Ohrievač vzduchu, 5) Ventilátor, 6) Zvlhčovač, 7) Svetlá, 8) Senzor: EC, pH a teplota vody, 9) Senzor: teplota a vlhkosť vzduchu, 10) Senzor intenzity osvetlenia, 11) Raspberry PI, 12) Senzor prietoku a ventil, 13) Nádoby s roztokmi, 14) Rezervoár, 15) Vzduchovací kameň, 16) Pumpa

Na obr.1 je vidieť množstvo použitých senzorických a akčných členov, ktoré by mali po použití vhodného algoritmu pre ich správu spôsobiť stabilný a plynulý chod systému. Pre reguláciu takéhoto dynamického systému bola vybraná fuzzy kognitívna mapa, kde hodnoty, namerané na senzorech hydroponického systému slúžia ako vstup do fuzzy kognitívnej mapy a jej výstup sa prenáša na akčné členy, ktoré ovplyvňujú hodnoty meraných veličín. Schéma fuzzy kognitívnej mapy použitej na riadenie takého systému sa nachádza na obr. 2.



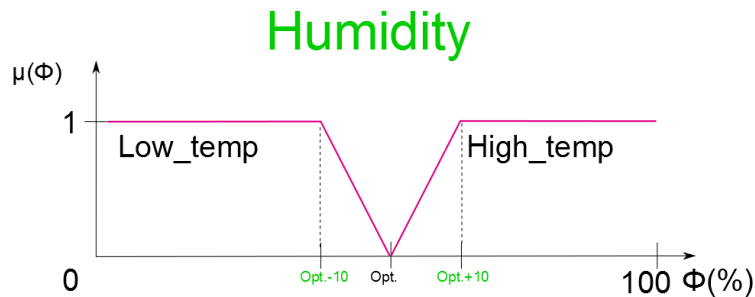
**Obr. 2 – Riadiaca fuzzy kognitívna mapa.**

Keďže sa jedná o fuzzy prístup, je vhodné vytvoriť pre všetky namerané hodnoty funkcie príslušnosti, či už pre ich nízku alebo vysokú hodnotu, tak ako napríklad pre hodnoty elektrickej vodivosti na obr.3.



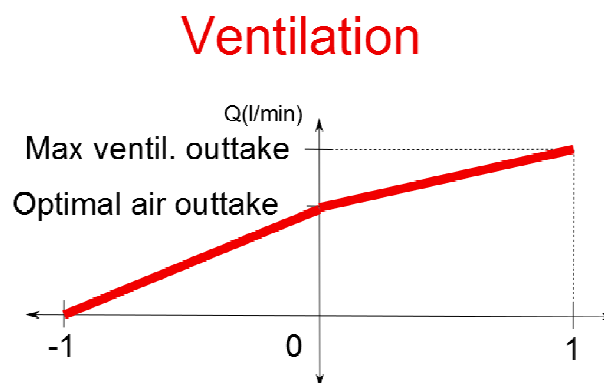
Obr. 3 – Funkcie príslušnosti elektrickej vodivosti.

Elektrická vodivosť ako aj ostatné faktory živného roztoku majú určitý vhodný rozsah na pestovanie plodín, v ktorom nie je nutné nijak zasahovať, a to napríklad z dôvodu oneskorenia odozvy živného roztoku na impulz. To však neplatí pre faktory vzduchu. Vzduch ako redšie médium reaguje na impulzy omnoho rýchlejšie a je omnoho jednoduchšie udržiavať jeho optimálne hodnoty pri nastavení funkcií príslušnosti, ako napríklad na obr. 4, kde sú zobrazené funkcie príslušnosti pre relatívnu vlhkosť.



Obr. 4 - Funkcie príslušnosti vlhkosti.

Výstup z fuzzy kognitívnej mapy je potrebné defuzzifikovať použitím transformačnej funkcie prenášajúcej hodnoty z intervalu  $< -1 ; 1 >$  do hodnôt nastavenia daného akčného člena v reálnej situácii. Na obr.5 je zobrazená transformačná funkcia určená pre nastavenie ventilácie.



Obr. 5 – Výstupná defuzzyfikačná funkcia pre ventiláciu

Na obrázku je vidieť, že v prípade, keď je na výstupe ventilácie nula, prítok vzduchu je rovný zhruba jednej pätine objemu miestnosti za minútu. Čím sa zabezpečí vhodná obnova vzduchu v miestnosti. V inom prípade prítok vzduchu lineárne rastie alebo klesá.

Celý systém je ovládaný pomocou Raspberry Pi, ktorý komunikuje s vnútornou databázou, kde je pod jednotlivými kľúčmi zapísaný aktuálny stav a vlastnosti systému. Táto databáza je



zosynchronizovaná s databázou na vzdialenom cloudovom servery, kde sa v prípade online módu vykonávajú všetky výpočtové kroky.

Samotná regulácia prostredia však nestačí, preto do aplikácie pribúda nová časť, ktorá ovplyvňuje rozvrhy svietenia a zalievania navrhnuté pomocou evolučného algoritmu. V prípade, že „Life Support“ zaznamená, že svetlá alebo vodné pumpy nepracovali dlhšie ako je vhodné tak sa automaticky zapnú a naopak.

Súčasťou riešenia algoritmu bolo vytvorenie a implementovanie postupu pre výmenu živného roztoku, ktorá by sa mala prevádzkať raz za isté obdobie (cca. 14 dní).

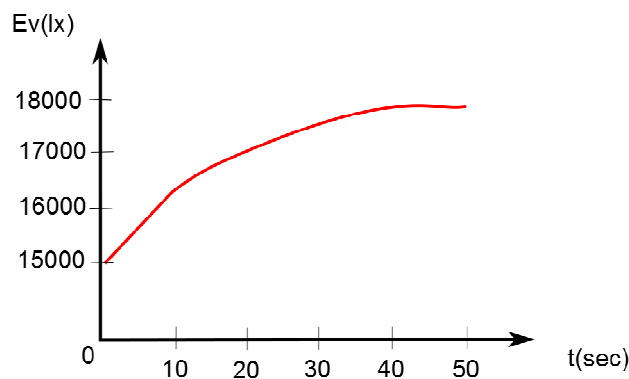
#### IV. EXPERIMENT

Nasledujúci experiment bol prevedený pre otestovanie reakcie systému na zaznamenaný pokles svetelnej intenzity, pričom požadovaná hodnota intenzity osvetlenia je 18000 luxov a tento experiment v čase nula začína na hodnote 15000 luxov. V priebehu dvoch sekúnd sa systém dostáva na optimálne hladiny svetelnej intenzity a neustále približuje jej hodnotu k prednastavenej, čo znamená že rastlinám je dodávané presne tolko svetla koľko naozaj v danom okamihu potrebujú. Nižšie sú uvedené grafy a tabuľka dokumentujúce priebeh experimentu.

t	Sensed Illuminance (lx)	Lights power (%)	FCM intervention (%)
0	15000	50	5,625
1	16400	55,625	3
2	17060	58,625	1,875
3	17500	60,5	0,498
4	17734	60,998	0,267
5	17858	61,265	

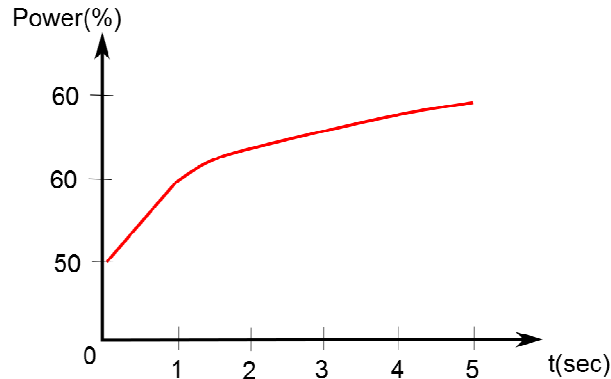
Tab. 1 – Tabuľka nameraných hodnôt

Na tabuľke č.1 je vidieť čas t v sekundách, ďalej snímanú intenzitu osvetlenia v luxoch, aktuálny výkon svetiel v percentách a nakoniec veľkosť vypočítanej zmeny výkonu svetiel. Z tabuľky je možné si predstaviť správanie sa systému a schopnosť sa prispôsobiť nečakanej zmene osvetlenia napríklad z dôvodu mrakov.



Obr. 6 – Priebeh intenzity osvetlenia v čase.

Graf znázorňuje priebeh nameranej intenzity osvetlenia v priebehu 5 sekundového experimentu.



Obr. 7 – Priebeh výkonu svetiel v čase

Graf predstavuje aktuálny výkon svetiel v priebehu experimentu.

#### V. ZÁVER

Táto práca prispieva k automatizácii hydroponického pestovania a k projektu G.E.C.K., ktorý sa zaoberá vývojom autonómnych hydroponických a aeroponických systémov. Ako už bolo načrtnuté v budúcnosti môže automatizácia takýchto procesov značne pomáhať pri získavaní nových zdrojov potravy. Do budúcnosti je potrebné pridať k práci reguláciu CO<sub>2</sub>. Práca obsahuje len popis základného systému, ktorý funguje správne pri experimentálnych podmienkach a pre absolútnu autonómnosť je potrebné mnohé upraviť.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### POUŽITÁ LITERATURA

- [1] TEXIER, W.: *Hydroponie pro každého*. Paris, Mama edition, 2013. p. 288. ISBN 978-2-84594-091-8
- [2] MLS, K.: *Fuzzy kognitívny mapy*, Dizertačná práca, Košice, TU-FEI, 2005.p. 86
- [3] RESH, H. M.: *Hydroponic food production*. USA, CRC Press, 2013. 489 s. ISBN 978-1-4398-7867-5
- [4] PAĽA, M.: *Navigácia autonómneho vozidla pomocou fuzzy kognitívnych máp*, Bakalárska práca, Košice, TU-FEI, 2009. p. 41.
- [5] PAĽA, M – MIZENKO, L - MACH M - REED T.: *Aeroponic Greenhouse as an Autonomous System using Intelligent Space for Agriculture Robotics*. Denver (USA). 2013
- [6] Global garden: Hystory of hydroponics [online]. dostupne na : <<http://www.rain.org/global-garden/hydroponics-history.html>> [cit. 2014-03-25]
- [7] BUGBEE, B.: *Nutrient managment in recirculating hydroponic culture*, Acta Hort, 2004. (ISHS) 648:99-112

# Analýza vložených páuz a dysfluentných javov v spontánnom rečovom prejave

Ján STAŠ, Marek PETRÁŠ, Jozef JUHÁR

Katedra elektroniky a multimediálnych telekomunikácií,  
Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach,  
Park Komenského 13, 041 20 Košice, Slovenská republika

jan.stas@tuke.sk, marek.petras@student.tuke.sk, jozef.juhar@tuke.sk

**Abstrakt** – Tento článok je zameraný na komplexnú analýzu vložených páuz a dysfluentných javov v anotovaných prepisoch rečových nahrávok v slovenskom jazyku obsahnutých v akustických databázach diktovaných súdnych rozhodnutí, spravodajských a diskusných relácií a parlamentných debát. Keďže v súčasnosti sú štatistické modely slovenského jazyka trénované z veľkej časti na textových korpusoch obsahujúcich písanú podobu jazyka, pri modelovaní spontánnej reči je potrebné mať k dispozícii okrem písaného textu aj reálne prepisy hovorenej reči obsahujúce miesta vzniku mimorečových a suprasegmentálnych javov pochádzajúcich priamo od rečníka. Analýza vložených páuz a dysfluentných javov je tak prvým krokom v procese zvyšovania kvality a robustnosti akustických a jazykových modelov v automatickom rozpoznávaní plynulej spontánnej reči v slovenčine. Na základe tejto analýzy je potom možné sa zamerať na najviac vyskytované mimorečové a suprasegmentálne javy v hovorenej reči a vytvoriť pre ne špecifický akustický a jazykový model, pomocou ktorého môžeme písaný jazyk čo možno najviac prispôbiť spontánnemu rečovému prejavu.

**Keywords** – automatické rozpoznávanie plynulej reči, dysfluentné javy, štatistické modelovanie jazyka, vložené pauzy

## I. ÚVOD

Pri tvorbe štatistických modelov jazyka použiteľných v systémoch na automatické rozpoznávanie plynulej reči (v angl. "*continuous speech recognition*", skr. CSR) je často nutné sa vysporiadať aj s rôznymi mimorečovými a suprasegmentálnymi javmi, ktoré pochádzajú priamo od rečníka a sú spôsobené najmä zlou výslovnosťou, nevhodnou artikuláciou, či nedokonalosťou rečového prejavu. Modely vložených páuz (v angl. "*filled pauses models*") a dysfluentných javov (v angl. "*models of disfluent speech*") sa preto snažia podchytiť a zahrnúť do jazykového modelu rôzne mimorečové a suprasegmentálne javy obsiahnuté v spontánnom rečovom prejave, ako napr. zaváhanie, zajakávanie, opakovanie fráz, skomolenie slov, či bežne vyskytujúce sa reflexy hovoriaceho, ako je hlboký nádych a výdych, mľaskanie, odkašľávanie si, či časté poťahovanie nosom. Vložené pauzy a dysfluentné javy častokrát vo vysokej miere vplývajú aj na celkovú chybovosť systému na automatické rozpoznávanie plynulej reči. Odhad relatívnej pravdepodobnosti týchto javov sa potom získava zväčša z anotovaných prepisov rečových nahrávok obsiahnutých v akustickej databáze a používaných v procese tréningu akustického modelu. Vhodným výberom správnej reprezentácie vybraných vložených páuz a dysfluentných javov v slovníku výslovnosti, akustickom modeli a štatistickom modeli jazyka, je potom možné dosiahnuť výrazné zlepšenie presnosti rozpoznávania reči, ako to bolo ukázané napr. v systéme na automatický prepis diktátu z oblasti súdnictva [1] zavedením relatívne malej množiny javov obsahujúcich skupinu vložených páuz, mimorečových prejavov rečníka a zvukov z okolia do modelu slovenského jazyka [2].

V tomto článku sú zhrnuté výsledky automatickej extrakcie vybraných mimorečových a suprasegmentálnych javov v anotovaných prepisoch rečových nahrávok v slovenskom jazyku z akustickej databázy obsahujúcej približne 5,75 mil. textových jednotiek (tokenov) v 680 hodinách plánovanej (imitovanej, diktovanej) a spontánnej reči. Tieto výsledky budú nápomocné pri tvorbe špecifických modelov vložených páuz a dysfluentných javov v akustickom a jazykovom modeli s cieľom zvýšiť kvalitu a robustnosť systémov na automatické rozpoznávanie spontánnej reči v slovenčine, tzn. systému na automatický prepis diktátu z oblasti súdnictva, na prepis spravodajských relácií, parlamentných debát a debát v malých konferenčných miestnostiach s viacerými účastníkmi.

## II. SÚČASNÝ STAV

Súčasný stav problematiky modelovania vložených páuz a dysfluentných javov v systémoch na rozpoznávanie reči sa v posledných rokoch opiera o poznatky z viacerých svetových výskumov.

A. Stolcke ako prvý predstavil jazykový model, ktorý určuje rozloženie vložených páuz, opakovaní a zmazaní s využitím pravdepodobnosti týchto javov v hovorenej reči. Výsledky modelovania však ukázali len minimálny rozdiel v znižovaní chybovosti systému na automatické rozpoznávanie reči, spôsobených pravdepodobne nízkou frekvenciou týchto javov v databáze diktovanej reči [3].

V práci [4] autori poukázali na odlišnú funkciu vložených páuz a opakovaní v závislosti od toho, na akom mieste v reči (vo vete) sa tieto javy vyskytujú. Pre vybranú skupinu opakovaní a vložených páuz vytvorili redukovanú množinu značiek pričom brali do úvahy aj okolitý kontext, v ktorom sa daný jav vyskytuje. Na základe tejto distribúcie vytvorili špecifický model, pomocou ktorého simulovali rozloženie týchto javov v písanom texte. Po začlenení vybraných značiek do tréningového procesu, výsledky perplexity jazykového modelu vykazovali iba mierne zlepšenie.

V príspevku [5] sa autori zamerali na identifikáciu miest vzniku modifikácií častí viet a opakovaní slov alebo celých viet. Po identifikácii týchto miest zostavili systém CSR, ktorý pozostával z akusticko-prozodického modelu, modelu obsahujúceho vložené pauzy a dysfluentné javy v jazyku a jazykového modelu založeného na triedach slov. Vhodnou kombináciou týchto modelov dospeli k značnému zlepšeniu systému ako celku pri rozpoznávaní spontánnej (konverzačnej) reči.

J. Duchateau a kol. na zlepšenie predikcie a robustnosti modelov jazyka využili okolitý kontext, ktorý predchádza vybranej skupine zaváhania a opakovaní. Bolo zistené, že úspešnosť rozpoznávania je možné zvýšiť, ak systém CSR má možnosť voľby, že daný jav, na základe jeho distribúcie v tréningových dátach, bude ponechaný alebo odstránený zo sekvencie dát po rozpoznaní [6].

Modely vložených páuz boli použité aj pri návrhu systému na automatické generovanie syntetickej reči z textu [7]. Autori sa zamerali na takú množinu mimorečových javov, ktoré sa početnosťou vyrovnávajú najviac frekventovanejším slovám v hovorenej reči. Na tréning vybranej skupiny vložených páuz využili priemerný model reči. Výsledky ukázali, že takýto model môže byť použitý aj pri generovaní vložených páuz bez narušenia kvality pôvodnej syntézy reči.

## III. ZÁKLADNÉ ROZDELENIE

Spontánna reč sa od diktovanej, resp. plánovanej reči líši množstvom mimorečových a suprasegmentálnych javov, ktoré narúšajú plynulý prejav rečníka, a vnášajú do samotného prejavu mnoho nejednoznačností. Tieto systém na automatické rozpoznávanie reči nedokáže správne vyhodnotiť, a dochádza tak k chybám. V rámci kategorizácie ich rozdelíme do troch základných skupín:

- **Izolované šumy.** Do tejto kategórie môžeme zaradiť najmä prejavy emócie ako smiech alebo plač, popevok, či pískanie si, kýchnutie, ale aj šepot a neartikulované zvuky, ktoré nemusia pochádzať priamo od rečníka, môžu byť aj produktom okolia. Tieto mimorečové prejavy nie je možné priamo predpovedať a častokrát nevieme odhadnúť ich miesta vzniku, ani dobu ich trvania. Z toho dôvodu, analýza izolovaných šumov nebola predmetom ďalšieho výskumu v oblasti zvyšovania kvality a robustnosti štatistických modelov slovenského jazyka.
- **Dysfluentné javy** sú reprezentované najmä poruchami reči, ktoré narúšajú prirodzený rečový prejav hovoriaceho. Častokrát za nimi nasleduje aj zmena pôvodnej myšlienky. Medzi dysfluentné javy radíme najmä *prolongácie*, rôzne druhy *repetícií*, ako opakovanie slov, slovných fráz alebo celých viet, modifikácie slov a slovných fráz vnútri viet, či patologické poruchy sprevádzané *zajakávaním*, resp. koktaním, ktoré rečník vkladá do spontánneho rečového prejavu nekontrolovane. Doba ich trvania sa odhaduje od niekoľkých milisekúnd až po jednotky sekúnd, v závislosti od toho, či došlo k predĺženiu, resp. opakovaniu hlásky, skupiny hlások, slov, slovných fráz, alebo celých viet. Vyskytujú sa najmä na začiatku viet, alebo v miestach dlhšej pauzy, kedy sa rečník v spontánnom prejave pozastaví a premýšľa, čo povie ďalej. Prípadne v stresových a vypätých situáciách, kedy rečník nedokáže kontrolovať svoj vlastný prehovor so zvyšujúcim sa tempom reči.
- **Vložené pauzy** môžeme rozdeliť do dvoch základných skupín, a to na zaváhania, resp. *hezitácie* a *reflexy*. Do skupiny hezitácií radíme všetky rečové a mimorečové prejavy zaváhania pochádzajúce priamo od rečníka. V anotáciách rečových nahrávok sa označujú najmä ako švy (ə), neartikulované hezitácie ([hez]), v prípade, ak ich vieme reprezentovať grafémovým zápisom, je možné použiť aj vlastnú triedu označujúcu citoslovce (%hmm, %mhm, %eh, a pod.). Na druhej strane, medzi reflexy patria všetky periodicky opakujúce sa javy, ktoré rečník vkladá do rečového prejavu najmä na začiatku a konci viet, alebo v prípade dlhších súvetí v miestach prestávky pre hlboký nádych alebo výdych. Z pohľadu akustického a jazykového modelovania patria medzi najfrekventovanejšie a najviac predpovedateľné javy.

Podrobnejšej analýze vybraných druhov vložených páuz a dysfluentných javov v anotovaných prepisoch rečových nahrávok v slovenčine sa budeme venovať vo výsledkovej časti tejto práce.

#### IV. ZDROJOVÉ DÁTA

K automatickej analýze vložených páuz a dysfluentných javov boli použité nasledovné akustické databázy obsahujúce anotované prepisy rečových nahrávok v slovenskom jazyku [1][8]:

- **APD** databáza obsahuje približne 236 hodín anotovaných rečových nahrávok diktovaných súdnych rozhodnutí, foneticky bohatých viet, novinových článkov, zaužívaných fráz a hláskovaných slov v rozsahu 1,86 mil. tokenov. Databáza je vyvážená čo sa pohlavia rečníkov týka a z veľkej časti obsahuje plánovanú diktovanú reč, s malým zastúpením dysfluentných javov.
- **BN** databáza pozostáva z anotovaných prepisov spravodajských relácií z ranných a večerných televíznych a rozhlasových novín v rozsahu približne 238 hodín a 2,47 mil. tokenov. Z troch štvrtín obsahuje diktovanú reč od profesionálnych rečníkov, zvyšná časť databázy je tvorená spontánnym, častokrát aj gramaticky nespisovným prejavom od veľkého množstva rečníkov.
- **PAR** databáza je tvorená prepismi parlamentných debát realizovaných v hlavnej rokovacej sále Národnej rady Slovenskej republiky v rozsahu približne 750 tis. tokenov obsiahnutých v 125 hodinách anotovaných prepisov rečových nahrávok. Databáza sa vyznačuje nevyváženým zastúpením pohlavia rečníkov a obsahuje polospontánny až spontánny prejav.
- **SUS** databáza pozostáva z 80 hodín (687 tis. tokenov) anotovaných prepisov diskusných relácií "Súdna sieň". Vyznačuje sa vysokou mierou spontánnosti prejavu, ktorý bol vytvorený amatérskymi a poloprofesionálnymi hercami.

Štatistika celkovej dĺžky trvania rečových databáz rozdelených podľa pohlavia rečníka a stupňa spontánnosti prejavu, ako aj celkového počtu rečníkov je zobrazená v Tab. 1.

Tabuľka 1  
 Akustická databáza.

akustická databáza		APD	BN	PAR	SUS
celkový	počet tokenov	1 860 699	2 470 061	750 050	687 075
	čas spolu	236:02.45	238:35.53	125:14.05	80:17.20
podľa stupňa spontánnosti	spontánna reč	32:15.40	63:30.17	42:06.51	55:03.19
	plánovaná reč	199:21.19	164:07.16	80:01.51	14:05.35
podľa pohlavia	muži	118:32.19	125:11.01	99:41.12	34:55.15
	ženy	114:34.32	110:11.34	22:32.20	27:46.37
počet rečníkov	muži	134	9 102	343	264
	ženy	137	4 998	98	218

Celková veľkosť akustickej databázy je 680 hodín, z toho približne 18 hodín je vyčlenených na testovanie systému na automatické rozpoznávanie plynulej reči ako celku a ďalších 12 hodín obsahuje segmenty reči, ktoré nie sú vhodné ani na tréningovanie akustického modelu, ani v procese jeho testovania. Akustická databáza obsahuje celkovo 457 hodín nahrávok plánovanej diktovanej reči a 193 hodín spontánneho rečového prejavu. Čo sa týka pohlavia rečníkov, ženský hlas je zastúpený v 274 hodinách, zvyšných 376 hodín sú rečové nahrávky mužského pohlavia.

Z pohľadu anotovaných prepisov rečových nahrávok, databáza obsahuje celkovo 5,75 mil. textových jednotiek, ktoré boli použité k automatickej extrakcii vložených páuz a dysfluentných javov pomocou nami vytvorených nástrojov v programovacom jazyku PERL.

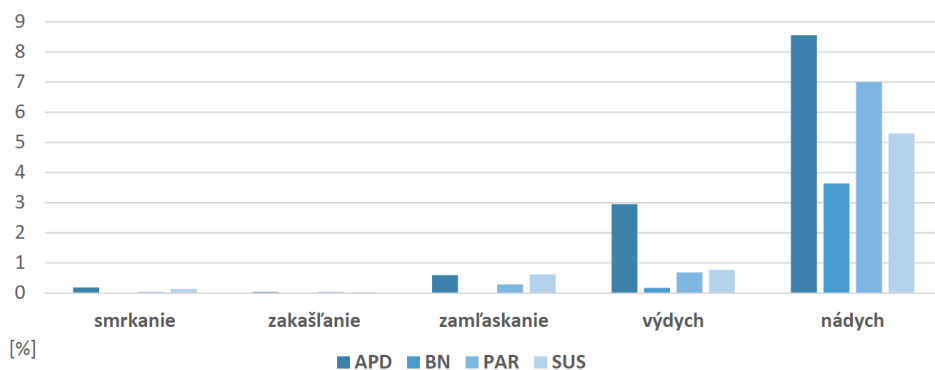
#### V. VÝSLEDKY

V tejto kapitole sa budeme venovať výsledkom automatickej extrakcie a analýze vybraných typov vložených páuz a dysfluentných javov v štyroch rečových databázach, ktoré boli predstavené v predchádzajúcej časti. Kvôli lepšej prehľadnosti sme ich rozčlenili do samostatných podkapitol.

##### A. Analýza vložených páuz

Pri analýze vložených páuz sme sa zamerali na dve skupiny javov, ktoré zaradzujeme medzi vložené pauzy, a to na: a) *reflexy*, ktoré patria do skupiny mimorečových prejavov rečníka b) *hezitácie*, ktoré sú zvyčajne sprevádzané krátkou rečovou aktivitou.

1) *Reflexy*: sú prirodzenou súčasťou prejavu každého rečníka a ako už bolo spomenuté, patria k mimorečovým prejavom s periodicky, alebo neperiodicky sa opakujúcou aktivitou hovoriaceho, ako je *nádyh* a *výdych*, v opačnom prípade môže ísť o *pokašľávanie*, *mlaskanie*, či *smrkanie*. Vzhľadom na to, že doba trvania reflexov je najkratšia spomedzi všetkých mimorečových a suprasegmentálnych javov, ich frekvencia je preto najvyššia a závislá najmä od pohody hovoriaceho a jeho zdravotného stavu. Na Obr. 1 je graficky znázornené percentuálne zastúpenie vybraných typov reflexov v dostupnej rečovej databáze. Z grafu možno vidieť, že najvyššie zastúpenie majú reflexy pre nádyh, ktorých priemerná hodnota dosiahla úroveň 6,12% spomedzi všetkých



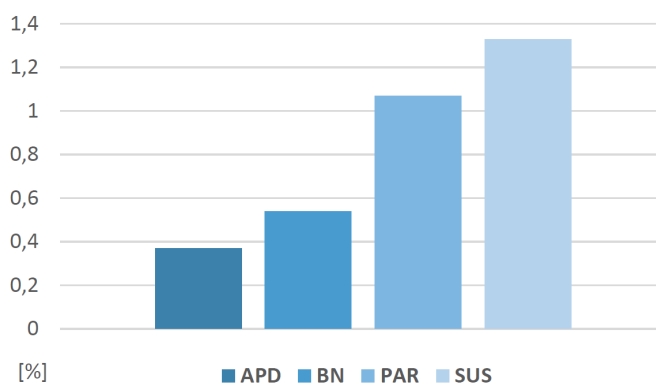
Obr. 1 Percentuálne zastúpenie vybraných typov reflexov v databáze prepisov rečovných nahrávok.

tokenov. Reflexy pre hlboký výdych dosahovali hodnotu priemerne 1,14% a pre zamľaskanie priemerne 0,38%. Najnižšie hodnoty sme zaznamenali pre posmrkávania a pokašľávanie. Ostatné mimorečové prejavy dosahovali zanedbateľné hodnoty, preto neboli zahrnuté do pozorovania.

Tabuľka 2  
 Štatistika vybraného počtu najviac vyskytovaných hezitácií v databáze prepisov rečovných nahrávok.

APD		BN		PAR		SUS	
ə	4 190	%eh	3 352	ə	7 373	ə	6 845
[hez]	1 244	ə	2 672	[hez]	250	%m	438
%m	130	%ehm	1 013	%m	114	%mhm	321
%mhm	96	[hez]	983	%mhm	22	[hez]	303
%hm	31	%eeh	607	%eh	5	%hm	164
%mm	11	%ehh	461	%em	5	%mm	115
%hmm	11	%hm	256	%mm	1	%mh	19
%ehm	9	%ah	216	%hm	1	%uhm	15
%ahm	9	%hmm	86	%mh	1	%hmm	6
%uhm	6	%m	73	%ah	1	%ehm	4

2) *Hezitácie*: patria rovnako ako reflexy do skupiny vložených páuz. Označujú miesta vzniku zaváhania, kedy sa rečník vnútri vety alebo na jej konci pozastaví, premýšľa čo povie ďalej a tento čas vyplní krátkou rečovou aktivitou. V prepisoch rečovných nahrávok sú hezitácie rozdelené do troch základných skupín, a to na: a) švy (ə); b) *neartikulované hezitácie* ([hez]); a c) *citulosovcia* pre zaváhania (%hmm, %mhm, %uh, a pod.). Doba ich trvania a frekvencia je porovnateľná s početnosťou bežných slov v jazyku. V Tab. 2 je možné vidieť štatistiku 10-tich najviac pozorovaných hezitácií v dostupných akustických databázach, z ktorej môžeme zostaviť reprezentatívnu skupinu, ktorá bude zaradená do rozpoznávacieho slovníka a použitá pri akustickom a jazykovom modelovaní v systéme CSR. Ďalej, na Obr. 2 je graficky znázornené percentuálne zastúpenie všetkých pozorovaných hezitácií. Z grafu možno pozorovať, že so zvyšujúcim sa stupňom spontánnosti rečového prejavu rastie aj početnosť výskytu hezitácií tvorených rečníkmi.



Obr. 2 Percentuálne zastúpenie hezitácií v dostupných akustických databázach.

## B. Analýza dysfluentných javov

Medzi dysfluentné javy radíme všetky rečové poruchy pochádzajúce priamo od rečníka, ktoré narušujú jeho plynulý prejav. V určitých prípadoch pozastávajú jeho sled myšlienok alebo dochádza k zmene pôvodnej myšlienky za inú. Všeobecne dysfluentné javy delíme do troch skupín, a to na: a) *prolongácie*, kedy dochádza k predĺženiu hlásky vnútri alebo na konci slova; b) *repetície*, kedy dochádza k opakovaniu slova, slovnej frázy, alebo celej vety, alebo k modifikácii slov vnútri viet; a c) *zajakávania*, pri opakovaní začiatkovej hlásky, resp. skupiny hlások na začiatku rečového prehovoru, alebo v miestach zaváhania rečníka, či pauzy v dlhších súvetiach.

1) *Prolongácie*: sa vyznačujú akustickým natiahnutím hlásky v slove, čím dôjde k jej predĺženiu. Rečník si v danom okamihu zrejme niečo uvedomil, po čom chce zmeniť svoj prehovor, alebo pochybuje nad vlastnou výpoveďou. Tým, že dôjde k predĺženiu hlásky v slove, rozpoznávač môže danú hypotézu vyhodnotiť ako ukončenie slova dlhou samohláskou, alebo ako spojenie dvoch nasledujúcich hlások za sebou, pričom môže dôjsť rovnako tak k zret'azeniu dvoch slov do jedného slova, ako aj k rozdeleniu hypotézy na dve slová. Štatistiku 10-tich najviac vyskytovaných prolongácií v rámci dostupných rečových databáz je možné vidieť v Tab. 3. Z tabuľky možno pozorovať, že medzi najviac vyskytované prolongácie radíme také slovné druhy ako predložky, spojky a častice, ukončené samohláskou v rozsahu jednej, nanajvýš dvoch slabík. Výskyt prolongácií je často náhodný, preto pri tvorbe štatistických modelov pre systémy CSR, je vhodné ich distribúciu prispôbiť obdobne ich náhodnému výskytu v databáze prepisov rečových nahrávok.

Tabuľka 3  
 Štatistika vybraného počtu najviac vyskytovaných prolongácií v databáze prepisov rečových nahrávok.

APD	BN	PAR	SUS
a: 140	a: 468	a: 509	a: 1 303
na: 81	že: 193	že: 350	že: 522
s: 55	na: 117	o: 186	no: 291
o: 54	je: 86	na: 138	ale: 151
v: 42	to: 66	je: 64	ja: 148
sa: 33	sa: 52	sa: 63	to: 126
že: 32	o: 50	ale: 59	sa: 102
do: 30	aby: 29	za: 49	na: 77
za: 30	v: 29	ja: 47	takže: 65
pre: 18	ale: 28	aby: 38	je: 60

2) *Repetície a zajakávania*: môžu zmeniť obsah prejavu hovoriaceho omnoho výraznejšie. Rozlišujeme: a) *reštart vety*, kedy rečník v krátkom časovom slede opúšťa aktuálnu myšlienku, pozmení svoju výpoveď a začína prehovor od nového začiatku; b) *modifikáciu frázy*, ktorá vzniká vnútri vety, kedy sa rečník poopraví v sekvencii  $n$  za sebou nasledujúcich slov, pozmení aktuálnu myšlienku, ale vetu nezačne vravieť odznova, pričom tento jav je často sprevádzaný tiež zaváhaním; c) *skrátene opravy*, ktoré často vznikajú v momente, kedy si rečník uvedomí chybu vo svojom prehovore, najmä v prípade zámény slova s rovnakým základom no odlišnou koncovkou; a d) *opakovacie slová*, pri opakovaní jedného a toho istého slova viackrát za sebou (častý prípad neohybných slovných druhov, ako sú zámená, predložky, spojky, a pod.). Špeciálnym prípadom je opakovanie začiatkovej hlásky alebo slabiky slova pri *zajakávaní sa*, ktoré je doprovoďným prejavom nervozity, úzkosti a strachu, alebo naopak agresivity, zvyšujúceho sa tempa reči, prípadne ide o vrozenú či patologickú poruchu hovoriaceho. Tab. 4 zobrazuje reálne príklady opakovania a zajakávania sa a v Tab. 5 je zhrnutá  $n$ -gramová štatistika najviac vyskytovaných opakovacích slov, extrahovaných z dostupných databáz prepisov rečových nahrávok. Výskyt repetícií tak vzrastá, rovnako ako v prípade hezitácií, so zvyšujúcim sa stupňom spontánnosti rečového prejavu.

## VI. ZÁVER

Cieľom tohto článku bolo kategorizovať a analyzovať vložené pauzy a dysfluentné javy v dostupných prepisoch rečových nahrávok v slovenskom jazyku. Keďže súčasné modely slovenského jazyka v systémoch na automatické rozpoznávanie plynulej reči sú trénované na písaných textoch, pri tvorbe systémov na automatický prepis spontánnej reči je nevyhnutné prispôbiť písaný text čo možno najviac hovorenej podobe jazyka. To je možné dosiahnuť zavedením modelov vložených páuz a dysfluentných javov rovnako tak do rozpoznávacieho slovníka, ako aj do akustického a jazykového modelu. V súčasnosti existuje rad nástrojov na distribúciu rečových a mimorečových prejavov skrz tréningový korpus obsiahnutých v sade programových nástrojov SRILM [9], ktoré slúžia na tréning a hodnotenie kvality jazykových modelov, a ktorých použitie je opísané v [10]. Takýto spôsob modelovania mimorečových a suprasegmentálnych javov vedie k výraznému zlepšeniu robustnosti systému na automatické rozpoznávanie plynulej reči ako celku.

Tabuľka 4

Príklady opakovania a zajakávania sa, extrahované z reálnych prepisov rečových nahrávok.

repetície	príklady
reštart vety	Trestný čin ohrozenia. Trestný čin vraždy bol... Tak zamestnávateľ. Povinnosť zamestnávateľa je... Všetky aktivity robíme tak. Robíme všetky aktivity tak, aby...
modifikácia frázy	... možno sú to tie, možno sú to tí mladí ľudia... ... máme informáciu, že bolo, že vláda vyčlenila... ... minulý týždeň boli spres %ehh, boli spojzdnené všetky linky...
skrátene opravy	... mestská časť nemala žiadne, žiaden signál od ministra... ... od jadrovej elektrárne, ktoré, ktorá je vzdialená... ... kde bojovali, bojovalo medzi sebou približne...
opakovacie slová	A je je je podľa mňa veľká škoda pre každého herca... Ste pre to nič neurobili, páni poslanci, ani ani ani táto vláda... Keď som videl a počul, že ako kričia tam, že že že malý nedýcha...
zajakávanie sa	Navrhovateľka nebude požadovať od odp od odporkyne náhradu... Jej nedostatok je dôvodom na za za zamietnutie návrhu... To to toto ne ne není není pravda...

Tabuľka 5

Štatistika vybraného počtu najviac vyskytovaných opakovacích slov v databáze prepisov rečových nahrávok.

APD	BN	PAR	SUS
na na 5	aj aj 23	aj aj 28	si si 22
od od 4	že že 22	a a 19	že že 10
do do 2	na na 20	že že 18	ja ja 8
to to 2	to to 19	ale ale 9	nie nie 7
dlho dlho 1	ako ako 12	ktoré ktoré 9	ánoáno 6
rade rade 1	to to to to 4	na na na 2	že že že 6
na na na 1	je je je 2	ako ako ako 1	to to to 2
s s s 1	že že že 2	za za za 1	ja ja ja 1
od od od od 1	do do do do 1	je je je je je 1	si si si 1
že že že že 1	v v v v v v 1	o o o o o o 1	ani ani ani ani 1

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Výskum modulov pre inteligentné robotické systémy (kód ITMS projektu: 26220220141) na základe podpory operačného programu Výskum a vývoj, financovaného z Európskeho fondu regionálneho rozvoja (100%).

## LITERATÚRA

- [1] M. Rusko, J. Juhár, M. Trmka, J. Staš, S. Darjaa, D. Hládek, R. Sabo, M. Pleva, M. Ritomský, and S. Ondáš, "Recent advances in the Slovak dictation system for judicial domain," in *Proc. of the 6th Language and Technology Conference: HLT as a Challenge for Computer Science and Linguistics, LTC 2013, Poznań, Poland*, 2013, pp. 555–560.
- [2] J. Staš, D. Hládek, and J. Juhár, "Recent advances in the statistical modeling of the Slovak language," in *Proc. of 56th International Symposium ELMAR'2014, Zadar, Croatia*, 2014, accepted for publication.
- [3] A. Stolcke and E. Shriberg, "Statistical language modeling for speech disfluencies," in *Proc. of ICASSP, Atlanta, GA, USA*, 1996, pp. 405–408.
- [4] M.-H. Siu and M. Ostendorf, "Modeling disfluencies in conversational speech," in *Proc. of ICSLP, Philadelphia, PA, USA*, 1996, pp. 386–389.
- [5] Y. Liu, E. Shriberg, and A. Stolcke, "Automatic disfluency identification in conversational speech using multiple knowledge sources," in *Proc. of EUROSPEECH, Geneva, Switzerland*, 2003, pp. 957–960.
- [6] J. Duchateau, T. Laureys, and P. Wambacq, "Adding robustness to language models for spontaneous speech recognition," in *Proc. COST278 and ISCA Tutorial and Research Workshop (ITRW) on Robustness Issues in Conversational Interaction, Norwich, UK*, 2004, pp. 1–4, paper 11.
- [7] J. Adell, A. Bonafonte, and D. Escudero, "Modelling filled pauses prosody to synthesize disfluent speech," in *Speech Prosody 2010, Chicago, IL, USA*, 2010, pp. 1–4.
- [8] M. Pleva and J. Juhár, "Building of broadcast news database for evaluation of the automated subtitling system," *Communications*, vol. 15, no. 2A, pp. 124–128, 2013.
- [9] A. Stolcke, "SRILM - An extensible language modeling toolkit," in *Proc. of ICSLP, Denver, CO, USA*, 2002, pp. 257–286.
- [10] R. Nunes and L. Neves, "Filled pauses modeling," L2F - Spoken Language System Laboratory, INESC ID Lisboa, Portugal, Tech. Rep., 2008.



# Identifikácia autorít v konverzačnom obsahu

<sup>1</sup>Slavomír BELAJ, <sup>2</sup>Kristína MACHOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>slavomir.belay.2@student.tuke.sk, <sup>2</sup>kristina.machova@tuke.sk

**Abstract** — Bakalárska práca sa popisuje určovanie autority všeobecne ale sústreďuje sa na automatickú identifikáciu autorít na základe dát získaných zo štruktúry a obsahu diskusného fóra. Práca navrhuje prístup o odhadu stupňa autority v rámci webovej konverzácie a spôsobu vystupovania prispievateľa k webovým diskusiám. Sústreďuje sa na rozličné možnosti jej kvantitatívneho ohodnotenia. V prvej sekcii sa nachádza teoretický rozbor problému určovania autority. Druhá sekcia obsahuje návrh riešenia a jeho implementáciu. Tretia kapitola opisuje realizované experimenty. V závere je ohodnotenie získaných výsledkov a zhrnutie celej práce.

**Keywords** — Identifikácia autorít, webová diskusia, konverzačný obsah, dolovanie v textoch

## I. ÚVOD

Webové diskusie sú takou internetovou obdobou zasadacej miestnosti, kde sa stretne skupina ľudí a horlivo diskutuje o nejakej problematike. Tu ale nastáva zlom medzi priamou konverzáciou a tou internetovou. Veľakrát prisudzujeme váhu slovám na základe toho čo vidíme a toho ako človek, ktorý niečo hovorí vystupuje. Pri internetovej diskusii to ale nie je možné, čiže obraz toho akým spôsobom môžeme určovať autoritu práve v tejto oblasti sa výrazne líši. Je problémom určovať autoritu vo webových diskusiách. Keďže k dispozícii sú iba také informácie ako napríklad:

- Príspevky užívateľa
- Počet príspevkov vo fóre
- Počet príspevkov v danej diskusii
- Počet odpovedí na príspevky iných užívateľov
- Použitý slovník

Hlavným problémom, ktorý bolo potrebné vyriešiť je zistiť akým spôsobom by sa dalo určiť autoritu z dostupných dát a snažiť sa tieto výsledky čo najviac priblížiť subjektívnym odhadom. Cieľom práce bolo navrhnúť riešenie tohto problému a vyvinúť aplikáciu, ktorá by bola schopná automaticky generovať stupeň autoritatívnosti konkrétneho prispievateľa.

## II. AUTORITA VO WEBOVÝCH DISKUSIÁCH

Na hodnotenie konkrétneho obsahu príspevkov užívateľa je možné porovnávať dĺžku príspevkov. Človek, ktorý vie čo chce povedať a rozumie sa danej tematike spravidla napíše dlhší príspevok ako človek, ktorý sa o niečom informuje alebo kritizuje. Ďalším faktorom je používanie otázok. Ľudia, ktorí si prídu na diskusné fórum po radu majú väčšinou menej skúseností v danej problematike. Jedným kritériom by takisto mohla byť skladba príspevku. Autoritatívnejší ľudia sa zvyknú vyjadrovať v ucelenejšej forme a dávajú si pozor aby aj príspevok, ktorý napíšu vyzeral dobre.

Počet príspevkov vo fóre zvyčajne vyjadruje ochotu autora zapájať sa do diskusii a zároveň to poukazuje na tendenciu presadzovať svoj názor a konzultovať o téme s ostatnými používateľmi. Spravidla sa menej autoritatívni ľudia niečo opýtajú prípadne raz, dvakrát sa zapoja a zvyšok diskusie len sledujú vývoj konverzácie medzi ostatnými. Čiže väčší počet príspevkov znamená vyššiu autoritu.

V danej diskusii sa počtom príspevkov hodnotí najmä to ako sa užívateľ vyzná v danej problematike a ochotu o nastolenej téme diskutovať. Autoritatívnejší používateľ sa snaží vždy poradiť ostatným a preto vyšší počet príspevkov opäť znamená vyššiu autoritu.

Takisto je vhodné brať do úvahy počet reakcií na príspevky iných používateľov. Ľudia na fóre, ktorí si prišli iba po radu alebo sa zapájajú do konverzácie skôr náhodne disponujú menšou autoritou.

Posledným uvažovaným kritériom je použitý slovník. Aj keď obsah komunikácie tvorí iba malú časť celku napriek tomu sa dá analyzovať použiť na odhad autoritatívnosti daného používateľa a to tak, že porovnáme slová použité v príspevku s nejakým slovníkom, obsahujúcim tzv. kontrolné slová. K určovaniu autorít je možné zaujať akýkoľvek postoj a preto sa postupy líšia. Autoritu je možné posudzovať na základe objektívnych kritérií ale aj subjektívnych pocitov.

### III. NÁVRH ODHADU AUTORITY

Problém vyhľadávania autorít v konverzačnom obsahu, konkrétne v tomto prípade dolovaním dát z webovej diskusie, je riešený prístupom, ktorý berie do úvahy nielen informácie z konverzačného obsahu a konverzačnej štruktúry, ale aj slovník takzvaných „autoritatívnych slov“. Taktiež bola využitá metóda strojového učenia, konkrétne neurónové siete. Pôvodne mali byť konverzačné texty čerpané z portálu youtube.com ale ukázalo sa, že je to veľký problém. Ako ďalší sme chceli použiť portál SME.sk kde by sa mohla brať do úvahy aj takzvaná „karma“ prispievateľa ako doplnkový nástroj pre určovanie autority. Napokon bol použitý obsah portálu pforum.sk.

Počas návrhu a implementácie riešenia problému analýzy autorít v konverzačnom obsahu bolo potrebné čeliť viacerým problémom. Ako prvý problém bol spôsob získavania mien prispievateľov, všetky ich príspevky v konkrétnej diskusii ako aj počet príspevkov celkovo vo fóre a počet reakcií užívateľa na ostatných. Tento problém bol riešený pomocou rozkladu zdrojového (html) kódu fóra. Ďalším problémom bolo určiť akým spôsobom efektívne riešiť autoritu na základe daných vstupov. Chceli sme zahrnúť čo najväčšie množstvo týchto vstupov. Problémom bolo aj nájsť spôsob akým by mohol používateľ v prípade nutnosti priamo upraviť výsledné hodnotenie programu, ale aby takýto zásah nebol nutný a program mohol správne pracovať aj bez toho. Posledným, možno trochu menším oproti ostatným, problémom bolo správny a prehľadný výpis výsledného hodnotenia.

### IV. POPIS IMPLEMENTÁCIE

#### A. Rozkladanie zdrojového kódu stránky

Výber stránky pforum.sk mal tú výhodu, že používa jednotnú formu zdrojového kódu. Následne bolo ale potrebné riešiť problém s tým, že počet príspevkov na stránke bol len pätnásť a ak mala daná diskusia napríklad 245 príspevkov bolo potrebné rozkladať a analyzovať kód cez viac url adries. Našťastie sa url adresa priamo v programe dala upravovať takým spôsobom, že v nej stačilo meniť jedno číslo a opakovane potom získať a rozložiť kód.

Na ukladanie diskusie bol použitý list objektov v jazyku C#. List obsahuje všetkých prispievateľov a každý z nich nesie informáciu o počte príspevkov v danej diskusii, príspevkov na celom fóre, počte reakcií na ostatných užívateľov a nakoniec texty jeho príspevkov v danej diskusii. Všetky tieto informácie sú pri analýze zdrojového kódu ukladané do tried a sú následne použité pri hodnotení.

#### B. Hodnotenie autority

Počas implementácie riešenia analýzy autorít sa ukázalo, že najefektívnejšie hodnotenie bude získané ak sa skĺbia viaceré hodnotenia a výsledkom bude ich priemer. Ako prvé boli hodnotené texty príspevkov a keďže text sa skladá zo slov, bolo potrebné určiť hodnotu každého slova. Výsledná hodnota textu je kombináciou jeho dĺžky a formy (použitie otáznikov a pod.). Hodnota slov bola použitá pri určovaní, ktoré slová majú byť pridané do slovníka a budú použité na hodnotenie ďalšej diskusie. Ďalším kritériom bol počet príspevkov v danej diskusii a počet príspevkov celkovo vo fóre. Posledným kritériom boli reakcie na ostatných používateľov. Pri každom kritériu boli zoradení všetci používatelia podľa konkrétneho kritéria (napríklad podľa počtu príspevkov) a potom som od prvého boli pridávané k ratingu body od 10 znižované o 10/počet prispievateľov. Pre príklad ak bolo 5 používateľov tak ten, ktorý mal najviac príspevkov dostal 10 bodov ďalší dostal  $10 - 10/5 = 8$  bodov a ostatní 6,4, a 2.

Takto boli ohodnotení vo všetkých kategóriách a výsledkom bol priemer za všetky kategórie. Čiže ak mal jeden používateľ v týchto kategóriách hodnotenia 10, 6, 4, 10 tak jeho výsledný rating bol 6. Takéto hodnotenie bol použitý z dvoch dôvodov. Jedným z nich je rovnomerný tlak každého kritéria na výsledné hodnotenie a druhým je to, aby sa výsledný rating pohyboval v rozmedzí 0-10.

### C. Slovník

Na pomocné hodnotenie autorít v diskusií boli použité tri typy slovníkov. Každý z nich pozostával zo súboru slov, ktoré sa následne kontrolovali so slovami použitými prispievateľmi. Prvý z nich bol slovník, vytvorený počas chodu aplikácie. Obsahoval slová, ktoré by v danom kontexte mohli zvýšiť autoritu prispievateľovi, ktorý ich použil. Slovník bol vytvorený manuálne na základe skúmania fóra a priebehu niektorých diskusií.

Ďalším slovníkom bol takzvaný „Kladný slovník“. Slúžil na ukladanie slov, ktoré, ak ich prispievateľ použil, zvyšovali hodnotu jeho príspevkov. Bol priamo prístupný pomocou programu a ukladal sa mimo behu programu, takže bol prístupný aj po opätovnom spustení programu. Rovnakým spôsobom fungoval „Záporný slovník“, ktorý naopak znižoval hodnotu príspevku, v prípade, že sa našla zhoda nejakého slova z tohto slovníka so slovami príspevku.

Tieto slovníky boli použité najmä na to, aby navodili u používateľa pocit, že program je schopný sa prispôbovať dátam, ktoré doň boli vložené. Pre príklad pri úprave slovníka počas vkladania desiatich diskusií sa pri jedenástej diskusií objavila iné výsledky ako použitím pôvodného slovníka, alebo bez použitia slovníka, či už kladného, záporného alebo toho pôvodného.

### D. Neurónová sieť

Použitie neurónovej siete malo mať podobný účel ako použitie slovníka. Teda dať užívateľovi programu možnosť priamo zasiahnuť do výsledného hodnotenia programu. Bola použitá neurónová sieť typu backprop – spätné šírenie chyby, ktorá mala jeden vstupný neurón, dve skryté vrstvy s piatimi neurónmi a jeden výstupný neurón. Vstupné hodnoty boli v rozsahu od 1 do 5.

Najväčším problémom bolo nájsť vhodné vstupné dáta, aby táto metóda síce zasiahla do výsledného hodnotenia ale zároveň aby nevychýlila výsledok priveľmi. Nakoniec boli použité na tréning náhodne vybrané slová priamo z diskusie a pomocou spätnej väzby od používateľa im bola priradená požadovaná hodnota. Ako vstup bola použitá hodnota slova vypočítaná aplikáciou a očakávaná výsledná hodnota predstavovala hodnotenie slova používateľom. Natrénovaná sieť bola použitá na vypočítanie hodnoty všetkých slov diskusie a takto ohodnotené slová sa podieľali na hodnotení príspevkov.

## V. TESTOVANIE APLIKÁCIE

Výsledky testov boli zobrazené v kontingenčnej tabuľky [2] a transformované na hodnoty presnosti a návratnosti. Stupnica hodnôt autority bola rozdelená na dve časti. Prvou boli hodnoty 0 – 5,9 a druhou hodnoty 6 – 10.

Keďže aplikácia má za úlohu rozoznávať autoritatívne pôsobiacich ľudí boli vytvorené kontingenčné tabuľky pre vyššie uvedené intervaly hodnôt na stupnici pre každú metódu samostatne. Za názor experta bol považovaný môj subjektívny odhad autority a pomocou kontingenčnej tabuľky boli zapísané výsledky. Pričom za kladnú odpoveď sa považovala hodnota autority v rozmedzí 6 – 10. Presnosť a návratnosť bola počítaná pomocou vzorcov:

$$\pi = \frac{A}{A+B} \quad \rho = \frac{A}{A+C}$$

Presnosť  $\pi$  charakterizuje počet správnych výsledkov pozitívnej klasifikácie do danej triedy k celkovému počtu výsledkov pozitívnej klasifikácie. Návratnosť  $\rho$  je počet správnych výsledkov pozitívnej klasifikácie do danej triedy k celkovému počtu skutočne pozitívnych príkladov.

Tabuľka 1  
 Kontingenčná tabuľka pre aplikáciu na odhad autority

	Expert – áno	Expert – nie
Program – áno	3	0
Program – nie	1	1

$$\pi = \frac{A}{A+B} = \frac{3}{3+0} = 1 \cong 100\% \quad \rho = \frac{A}{A+C} = \frac{3}{3+1} = 0,75 \cong 75\%$$

Tabuľka 2  
 Kontingenčná tabuľka pre odhad autority aplikácie s pôvodným slovníkom

	Expert – áno	Expert – nie
Program – áno	3	0
Program – nie	1	1

$$\pi = \frac{A}{A+B} = \frac{3}{3+0} = 1 \cong 100\% \quad \rho = \frac{A}{A+C} = \frac{3}{3+1} = 0,75 \cong 75\%$$

Tabuľka 3  
 Kontingenčná tabuľka pre odhad autority aplikácie s tvorením slovníka

	Expert – áno	Expert – nie
Program – áno	4	0
Program – nie	0	1

$$\pi = \frac{A}{A+B} = \frac{4}{4+0} = 1 \cong 100\% \quad \rho = \frac{A}{A+C} = \frac{4}{4+0} = 1 \cong 100\%$$

Tabuľka 4  
 Kontingenčná tabuľka pre odhad autority aplikácie so slovníkom tvoreným pomocou NS

	Expert – áno	Expert – nie
Program – áno	3	1
Program – nie	1	0

$$\pi = \frac{A}{A+B} = \frac{3}{3+1} = 0,75 \cong 75\% \quad \rho = \frac{A}{A+C} = \frac{3}{3+1} = 0,75 \cong 75\%$$

Tabuľka 5  
 Porovnanie výsledkov z kontingenčných tabuliek

Metóda	$\pi$ - Presnosť	$\rho$ - Návratnosť
Samotný program	100%	75%
Pôvodný slovník	100%	75%
Slovník automaticky	100%	100%
Slovník pomocou NS	75%	75%

## VI. ZÁVER

Keďže nosnou myšlienkou tejto práce je určovanie autorít bolo potrebné v prvom rade skúmať čo autorita vlastne je a akým spôsobom sa prejavuje v technologickom svete so zameraním sa na webové diskusie. Na základe týchto vedomostí potom bolo potrebné nájsť vhodné diskusné fórum a z neho čerpať vstupné dáta. Výstupom bola aplikácia, ktorá bola schopná určiť autoritu v rámci zadanej diskusie. Túto úlohu sa podarilo splniť s primeranou presnosťou, keďže ide o novú, dosiaľ nepreskúmanú oblasť.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## POUŽITÁ LITERATÚRA

- [1] Sinčák, Peter – Andrejková Gabriela: Neurónové siete Inžiniersky prístup (1. diel) [online]. Košice: TU, FEI, 1996. [cit. 2014-05-14]. Dostupné na internete: <<http://www.ai-cit.sk/cigOld/source/publications/books/NS1/html/index.html>>. ISBN 80-88786-38-X
- [2] Machová Kristína: Strojové učenie v systémoch spracovania informácií [online]. Košice: TU, FEI, 2010. [cit. 2014-05-28]. Dostupné na internete: <<http://people.tuke.sk/kristina.machova/pdf/monoSUvSSL.pdf>>. ISBN 978-80-8086-130-8

# Limitations of Dispersion and Pulse Broadening in Optical Fiber

<sup>1</sup>Ján TÓTH, <sup>1</sup>Luboš OVSEŇÍK, <sup>1</sup>Ján TURÁN

<sup>1</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>jan.toth@tuke.sk, <sup>1</sup>lubos.ovsenik@tuke.sk, <sup>1</sup>jan.turan@tuke.sk

**Abstract** — An optical communication technology was studied and developed significantly in last few decades. Optical fiber is an excellent dielectric medium which is capable to carry a huge informational stream. Within an optical fiber there are basically two groups of negative effects which are needed to be studied and eliminated. The first group is called chromatic dispersion which will be taken under the scope in this paper. On the other hand there is the second group, generally called nonlinear effects. If we want to achieve a satisfactory data bit rate we have to keep balance between these two negative contributors. However a dispersion phenomenon has been discussed in numerous papers there is still a demand to explain and clearly understand this key topic.

**Keywords** — chromatic dispersion, material dispersion, waveguide dispersion, wavelength

## I. INTRODUCTION

There are certain problems which are needed to be solved in optical communication systems. Besides many others there is a dispersion playing role. Dispersion causes pulse broadening in time domain which leads to a higher bit error rate (incorrect detection). Dispersion is basically because a light has various frequency components which are travelling with different velocities. Dispersion has various forms however the most significant is so called chromatic dispersion. Chromatic dispersion is understood as a sum of material and waveguide dispersion. If one wants to calculate material dispersion the waveguide has to be neglected. On the other hand a calculation of waveguide dispersion can be made under the assumption that material dispersion is not playing role. In general dispersion is a negative effect however we are taking advantage of it in optical communication systems. Dispersion is compensated by so called nonlinear effects especially self-phase modulation phenomenon [1].

## II. DISPERSION LIMITATIONS

### A. Dispersion overview

Dispersion is one of the negative effects which are present in optical fiber when signal propagates through. When we look deeper into the topic, the dispersion itself is a quite complicated and complex phenomenon. In its simplest form dispersion causes a pulse broadening which is directly related to the bandwidth. It is obvious that a bandwidth creates one of the major constraints on data bit rate. Data bit rate and related indicators as SNR and BER are basically the most important factor which evaluate the optical link, its efficiency and reliability. If we want to understand how all the phenomena's work out together namely dispersion, a bandwidth and a data bit rate there is a high demand on mathematical comprehension among them [1]. The following equations will give us an explicit explanation, some basic ideas and problems which are needed to be solved and have been under the scope of researches for many years (1).

$$\Delta T = \frac{L}{c} \frac{n_1(n_1 - n_2)}{n_2} \quad (1)$$

where  $\Delta T$  is so called a pulse broadening parameter given in seconds,  $n_1$  and  $n_2$  are refractive indexes of core and cladding and  $c$  is the light velocity. From the ray light theory we understand that light beam is composed of some number of rays which represent different frequencies. Each ray travels with a particular velocity. In other words, the parameter  $\Delta T$  tells us about the time

difference between the slowest and the fastest ray. If there would be an ideal situation and we would have the only one ray or frequency respectively, there would be no pulse broadening present within the optical fiber. In a network transmission terminology we are using a parameter bandwidth  $B$  quite often (2). This parameter basically determinates how much data or information we are able to transfer through the network [2].

$$B = \frac{1}{\Delta T} \quad (2)$$

#### B. LED as an optical sources

Basic characteristics of LED as an optical source:

- typical spectral width is relatively high 30 nm,
- high dispersion (pulse broadening) and associated low data bit rate,
- extremely low efficiency,
- cannot be used in long-distance application,
- proportionality between optical power and LED current (current in junction),
- linear device –suitable for analog modulation (short distance applications).

#### C. Laser Diode as an optical source

Basic characteristics of laser diode as an optical source:

- high output power,
- high latching efficiency,
- spectral width 1-2nm,
- coherent communication (spatial and temporal), circuit compatibility
- ASK,FSK,PSK modulation formats supported,
- applicable in non-linear applications.

#### D. Optical sources

Talking about dispersion and its various forms we especially have to realize one more important aspect of this problematic. We are using gas lasers, LED and laser diodes as sources of light in optical links. An ideal solution for an optical source would be LASER application however there is a problem with compatibility with today's circuits. Speaking about all possible kinds of sources there is one key parameter playing role. It is so called spectral width  $\sigma_\lambda$  usually given in nano-meters. From the previous comparisons we're assuming that laser diode (especially DFB laser) is the most suitable choice in terms of optical sources because of their suitable  $\sigma_\lambda$ . One has to keep in mind that even if we are using laser diode as optical sources because of its narrow spectral width (approximately 1nm ~ 100 GHz). However this bandwidth is still very wide. It is important to note that the one optical source has multiple carriers. Each carrier is than modulated by the modulating signal and we can see the sidebands around each of the carrier frequency. So a broadening phenomenon is present because of different carriers which are modulated and travel with different velocities [1]. A bandwidth is then decided by the carrier which depends on optical source.

#### E. Dispersion Calculation

To calculate a mathematical expression for dispersion we have to define some basic parameters which are fundamental for optical communication systems. A *Group Velocity*  $v_g$  is basically a change in frequency with respect to propagation constant  $\beta$  given by the following (3). This will give us an idea which velocity is electric field moving with.

$$\omega = \frac{2\pi c}{\lambda} \quad v_g = \frac{\partial \omega}{\partial \beta} = 2\pi c \frac{\partial}{\partial \beta} \left( \frac{1}{\lambda} \right) \quad (3)$$

The second important parameter is so called *Group Delay* given in units of length. This parameter refers to value which is one upon  $v_g$ , represented by (4).

$$t_g = \frac{1}{v_g} \quad (4)$$

*Pulse Broadening*  $\tau_g$  is the next parameter which has to be taken into the account. Equation (5) shows us that spectral width  $\sigma_\lambda$  and group delay  $v_g$  with respect to the wavelength will finally decide about how much a pulse will be broadening.

$$\tau_g = \frac{dt_g}{d\lambda} \sigma_\lambda \quad (5)$$

By assuming and understanding all mentioned parameters we can define general expression which quantitatively evaluates dispersion and it is given by the (6).

$$D = \frac{dt_g}{d\lambda} \text{ [ps/km/nm]} \quad (6)$$

#### F. Dispersion forms

If one wants to understand the dispersion there is a very important fact which has to be considered. As we know we are using two main kinds of optical fibers in optical communication systems, single and multi-mode respectively. From the wave theory we know that the light is propagating in form of modes. That's why we are having three forms of dispersion:

- material dispersion  $D_{Mat}$
- intramodal dispersion  $D_{wg}$ ,
- intermodal dispersion.

### III. MATERIAL DISPERSION

Material dispersion (7) is basically given by the material properties. In other words we can say that once the material for an optical fiber is decided a material dispersion is determinate. This kind of dispersion is because of different frequencies travel with different velocities. And different velocities have some specific values for various materials. In fact we cannot influence material dispersion because it's fixed for a certain material.

$$D_{Mat} = -\frac{\lambda}{c} \frac{d^2 n(\lambda)}{d\lambda^2} \quad (7)$$

Experimentally was founded out that a material dispersion has minimal value around 1310 nm wavelength which is one of the so called optical windows. However zero dispersion is desirable there are certain reasons why we have to keep some amount of dispersion present because of nonlinear effects. From the practical analyses one can understand that we have positive (normal) and negative (anomalous) dispersion [4]. The difference between these two kinds of dispersion is shown in Fig. 1.

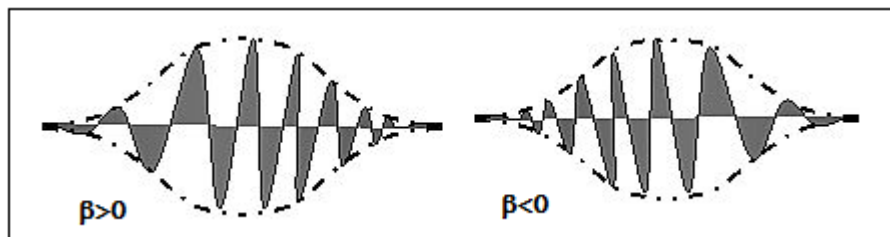


Fig. 1 Normal and anomalous dispersion

The numerical form for material dispersion is given by Equation 7. Material dispersion can be evaluated under the condition that refractive index is a function of wavelength and even more an influence of waveguide dispersion is neglected. Equation (8) shows us how the refractive index depends on wavelength.

$$n(\lambda) = C_0 + C_1 \lambda^2 + C_2 \lambda^4 + \frac{C_3}{(\lambda^2 - l)} + \frac{C_4}{(\lambda^2 - l)^2} + \frac{C_5}{(\lambda^2 - l)^3} \quad (8)$$

Where  $C_0=1.45$ ,  $C_1=-0.003$ ,  $C_2=-0.0000381$ ,  $C_3=0.0030$ ,  $C_4=-0.0000779$ ,  $C_5=0.0000078$  and  $l_0 = 0.035$ . A wavelength  $\lambda$  from this equation is given in  $\mu m$ . A graphical interpretation of material dispersion is shown in Fig. 2 [4].

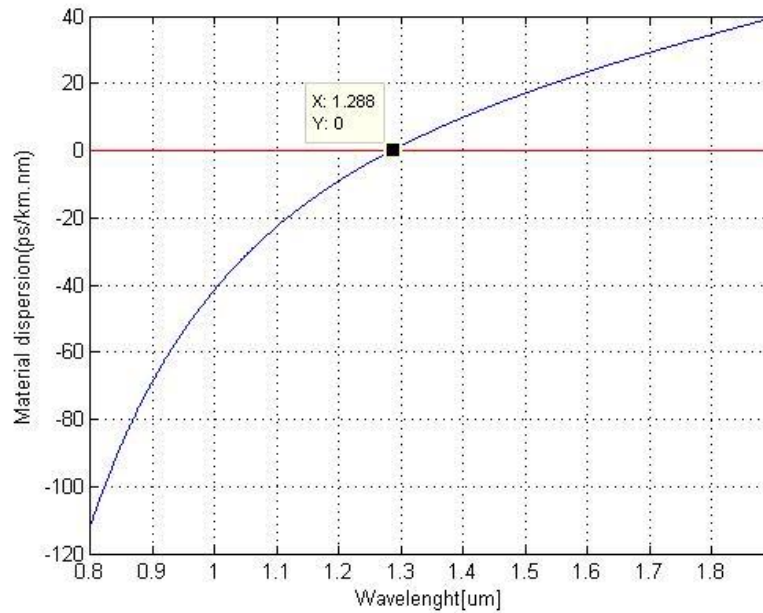


Fig. 2 Material dispersion

#### IV. INTERMODAL DISPERSION

##### A. Intermodal dispersion

Intramodal dispersion (often called waveguide) is a partly tunable parameter. In other words by modifying of some particular parameters of an optical fiber we can set up a desired dispersion value. Light is propagating through an optical fiber mostly in core however some amount travels in cladding. It is well known that refractive indexes for core and cladding are different and that's why the velocities in these two environments are also different. In fact the light travels faster in core and slower in cladding. The power is distributed in both core and cladding in some ratio. It is important to note that power distribution is wavelength dependent. Different power distributions cause that different wavelengths travel at different speeds. By modifying fiber's refractive index profile it is possible [6] to achieve desired waveguide dispersion. Waveguide dispersion is basically given by (9).

$$D_{wg} = -\frac{n_2 \Delta}{c \lambda} V \frac{d^2(bV)}{dV^2} \quad [ps / km.nm] \quad (9)$$

Waveguide dispersion formula depends on refractive indexes (core, cladding), relative refractive index  $\Delta$  which is basically a combination of  $n_1$  and  $n_2$ . A light velocity in Equation 9 is represented by  $c$  and  $V$  is a normalized frequency given by (10).

$$V = \frac{\omega}{c} a (n_1^2 - n_2^2)^{1/2} \quad (10)$$

$$V \frac{d^2(bV)}{dV^2} \cong 0.08.0 + 0.549.(2.834 - V)^2 \quad (11)$$

If we want to plot waveguide dispersion [5] as a function of wavelength we could use a certain approximation in (11). By using such an approximation we can get satisfying results for waveguide dispersion. A graphical interpretation of waveguide dispersion for fiber with core diameter of  $a = 9 \mu m$  is displayed in Fig. 3.



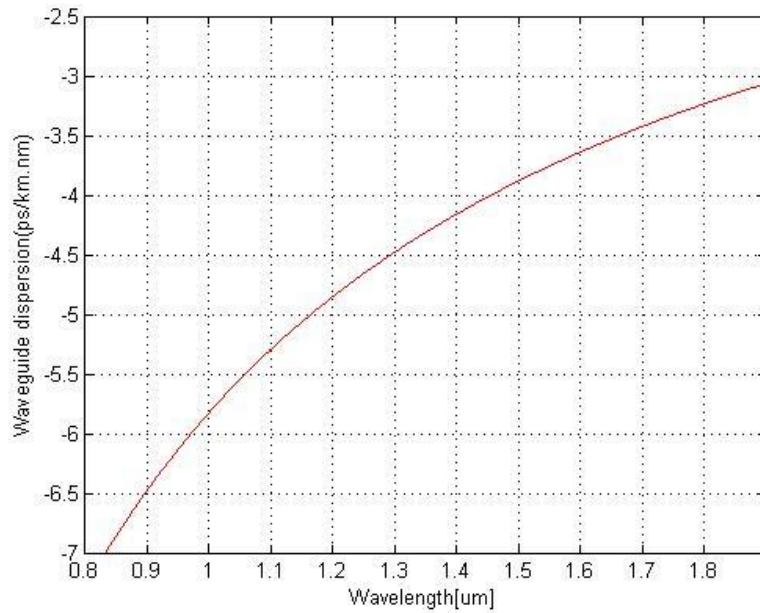


Fig. 3 Waveguide dispersion

### V. CHROMATIC DISPERSION

As it was already mentioned above that dispersion is a negative and not desired effect which is necessarily present when light propagates through the fiber. It is because of a light nature itself (different frequency components). Dispersion causes pulse broadening in time domain which leads to higher bit error rate and incorrect information interpretation on a receiver side. This effect is often called chromatic dispersion [7]. Chromatic dispersion is a sum of material and wavelength dispersions (12).

$$D_{CH} = \left( -\frac{\lambda}{c} \frac{d^2 n(\lambda)}{d\lambda^2} \right) + \left( -\frac{n_2 \Delta}{c \lambda} V \frac{d^2 (bV)}{dV^2} \right) \quad (12)$$

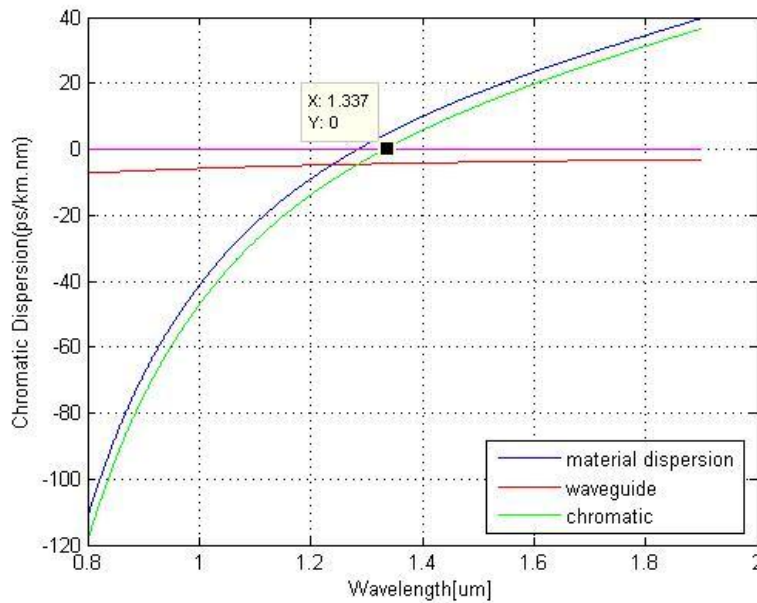


Fig. 4 Chromatic dispersion

## VI. CONCLUSION

The curvature of the chromatic dispersion is illustrated in Fig. 4. Chromatic dispersion is plot by red color and this curve crosses zero value at wavelength of 1.337  $\mu\text{m}$ . This value is often called a zero dispersion wavelength. The results from the Fig. 4 are corresponding with experimentally measured values of the chromatic dispersion. There can be small differences in a zero dispersion wavelength value among different literature sources however for some specific fiber attributes (refractive index of core and cladding, a core radius) the curve looks like shown in Fig. 4. It is very important to understand dispersion phenomena. It is not only because this is a negative effect in optical fiber which is needed to be eliminated, moreover we are using dispersion to suppress so called nonlinear effect (especially SPM – Self-Phase Modulation).

## ACKNOWLEDGMENT

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „063TUKE-4/2013 - The Use of Remote Controlled Optical Fibre Refractometer in Teaching“, and project no. „006TUKE-4/2014 - The Use of TUKE PON Experimental Model in Teaching“. This work is also the result of the contract no. “APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks” supported by the Slovak Research and Development Agency.

## REFERENCES

- [1] A. Govind, *Nonlinear Fiber Optics*. Inc. Waltham, MA 02451, Fifth edition 2013.
- [2] B. Chomycz, *Planning fiber Optic Networks*. Inc. McGraw-Hill Professional, 2009.
- [3] Y. Pochi, *Introduction to Photorefractive Nonlinear Optics*. Inc. New York, NY: John Wiley & Sons, 1993.
- [4] S. K. Raghuvanshi, S. Kumar, “Analytical Expression For Dispersion Properties Of Circular Core Dielectric Waveguide Without Computing  $\delta^2\beta/\delta\kappa^2$  Numerically,” in *I-manager’s Journal o Future Engineering & Technology*, 1 n Vol. 7 No. 13, 2012.
- [5] L. N. Binh, *Guided Wave Photonics: Fundamentals and Applications with MATLAB*. Boca Raton, FL: CRC Press Taylor & Francis Group, 2012.
- [6] H. Abramczyk, *Dispersion phenomena in optical fibers*. Lodz, PL.
- [7] M. Cvijetic, *Coherent and Nonlinear Lightwave Communications*. Artech House, Inc. Norwood, MA 02062, 1996.

# Chránenie malých energetických zdrojov

<sup>1</sup>Pavol VAŠKO, <sup>2</sup>Marek HVIZDOŠ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>pavol.vasko@student.tuke.sk, <sup>2</sup>marek.hvizdos@tuke.sk

**Abstrakt** — Príspevok je zameraný najmä na činnosť mini elektrární, pričom najväčšia pozornosť je venovaná podmienkam pre ich pripojenie do siete. K hlavným podmienkam pre pripojenie patrí správne nastavenie ochranných funkcií. Pritom je potrebné vychádzať zo skratových pomerov v sieti. Výsledkom je návrh a následné porovnanie chránenia malých zdrojov elektrickej energie.

**Kľúčové slová** — elektrické ochrany, malé zdroje energie, distribučná sieť, podmienky pripojenia

## I. ÚVOD

Do sústavy prevádzkovateľa distribučnej sústavy (PDS) je možné pripojiť len také zariadenie, ktoré svojim spätným pôsobením neovplyvňuje distribučnú sústavu a jej používateľov prekročením dovolených hodnôt parametrov elektrickej energie. Spôsob a miesto pripojenia určí prevádzkovateľ distribučnej sústavy s prihliadnutím k daným sieťovým pomerom, druhu, výkonu a spôsobu prevádzky zdroja, ako aj k oprávneným záujmom prevádzkovateľa zdroja. Tým ma byť zaručené, že zdroj bude prevádzkovaný bez rušivých účinkov a neohrozí ostatné zariadenia pripojené do distribučnej sústavy.

Pri výstavbe, rekonštrukcii a prevádzke elektrických zariadení je potrebné dodržiavať:

- platné zákony a vykonávacie predpisy,
- platné normy EN, STN, PNE prípadne PN PDS,
- predpisy pre ochranu pracovníkov a bezpečnosť práce,
- Technické podmienky PDS, Prevádzkový poriadok PDS, nariadenia a smernice PDS,
- Dispečerský poriadok.

Aby bolo zaistené dostatočné dimenzovanie zariadení, musí byť v každom prípade uskutočnený výpočet skratových pomerov v mieste pripojenia a zohľadnený skratový príspevok samotného zdroja. Skratová odolnosť zariadenia musí byť vyššia, nanajvýš rovná najväčšiemu vypočítanému celkovému skratovému prúdu.

## II. SPÍNACIE ZARIADENIA

Každý zdroj musí z hľadiska záujmov prevádzkovateľa distribučnej sústavy a bezpečnej a spoľahlivej prevádzky distribučnej sústavy obsahovať predovšetkým nasledovné spínacie zariadenia:

- rozpojovacie miesto (RM) – trvale prístupné spínacie zariadenie,
- hlavné rozpojovacie miesto (HRM) – väzobný spínač.

Rozpojovacie miesto predstavuje verejne prístupné spínacie miesto s funkciou odpájania za účelom viditeľného odpojenia zdroja od distribučnej sústavy (DS). V nízkonapäťových sieťach môže ako RM slúžiť prípojková skriňa alebo rozvodná istiacia skriňa za podmienok, že sú trvalo prístupné bez obmedzenia. V rozvážači transformátorovej stanice VN/NN, ak je do tohto vývodu vyvedený len jeden zdroj, je možné za trvale prístupné spínacie zariadenie taktiež považovať aj poistku alebo poistkový odpínač nízkeho napätia. Istič nespĺňa podmienky viditeľného rozpojenia preto ho nemožno použiť ako RM.

Vo vysokonapäťových sieťach je možné za RM považovať úsekový odpínač alebo odpájač v majetku prevádzkovateľa distribučnej sústavy, alebo v cudzom majetku, s podmienkou, že je trvalo prístupný pracovníkom PDS.

Pre zdroje s viacerými generátormi musí byť použité spoločne jedno trvale prístupné spínacie zariadenie. Spínač na spojenie zdroja s distribučnou sústavou slúži ako trvale prístupné spínacie miesto. Usporiadanie spínačov je závislé na pripojení, vlastníckych i prevádzkových pomeroch v odovzdávacej stanici. Bližšie údaje stanoví prevádzkovateľ distribučnej sústavy predpísaním technických podmienok a obchodných podmienok. RM má byť vo vlastníctve prevádzkovateľa distribučnej sústavy [1].

Pre spojenie zdroja s distribučnou sústavou je potrebné použiť väzobný spínač, ktorý má minimálne schopnosť vypínania záťaže a je mu priradená skratová ochrana. Tento väzobný spínač (hlavné rozpojovacie miesto) môže byť na nízkonapäťovej strane, ako aj na vysokonapäťovej strane. Ak sa nepredpokladá ostrovná prevádzka, je možné za tento prvok považovať spínacie zariadenie generátora. Spínacie zariadenie musí zabezpečiť galvanické oddelenie vo všetkých troch fázach. V prípade požiadavky prevádzkovateľa distribučnej sústavy má byť HRM alebo iné rozpojovacie miesto diaľkovo ovládané. Prevádzkovateľ distribučnej sústavy má právo určiť, ktoré spínacie prvky budú zaplombované. V prípade viacerých generátorov jedného zdroja sa použije jedno spoločné HRM.

Výpadok pomocného napätia pre ochrany a spínacie prístroje musí viesť automaticky k vypnutiu zdroja. Pre zdroje, kde je použitý striedač, je potrebné umiestniť spínacie zariadenie na striedavej strane striedača. Pri spoločnom umiestnení spínacieho zariadenia v skrini striedača, nesmie dôjsť k pôsobeniu spínača pri vzniknutí skratu v striedači. Vypnutím HRM by nemala byť obmedzená vlastná spotreba zdroja [2].

### III. SIEŤOVÉ OCHRANY A OCHRANY ZDROJOV

Aby funkčnosť distribučnej sústavy bola spoľahlivá a bezpečná, je potrebná dohoda medzi prevádzkovateľom zdroja a prevádzkovateľom distribučnej sústavy ohľadom systému chránenia. V tejto dohode sa prevádzkovatelia dohodnú na druhoch ochrán, vypínacích časoch, selektivitě a citlivosti sieťových ochrán. U zdrojov schopných ostrovej prevádzky je nutné zabezpečiť chránenie aj v ostrovej prevádzke.

Všeobecne je potrebné použiť ochrany s funkciami podľa Tab. 1.

Tab. 1 Nastavenie sieťových ochrán zdrojov

Nastavenie ochrán		Zdroje nad 100 kW		Zdroje do 100 kW	
Funkcia	Označenie	Max. hodnota	t (s)	Max. hodnota	t (s)
Nadpätie 1. st.	U>	110 %U <sub>f</sub>	0.5	110 %U <sub>f</sub>	0.1
Nadpätie 2. st.	U>>	120 %U <sub>f</sub>	0.1	—	—
Podpätie 1.st.	U<	90 %U <sub>f</sub>	0.5	85 %U <sub>f</sub>	0.1
Podpätie 2.st.	U<<	70 %U <sub>f</sub>	0.1	—	—
Nadfrekvencia	f>	51,5 Hz	0.1	51,5 Hz	0.1
Podfrekvencia	f<	47,5 Hz	0.1	47,5 Hz	0.1

V niektorých prípadoch môže byť s ohľadom na sieťové pomery potrebné iné nastavenie ochrán. Preto je ich nastavenie potrebné vždy dať odsúhlasiť prevádzkovateľovi distribučnej sústavy.

Vo štvorvodičových sieťach NN je nastavenie napätia vzťahované na napätie medzi fázami a stredným vodičom (fázové napätie). V trojvodičových sieťach VN a VVN je nastavenie napätia vzťahované na združené napätie medzi fázami.

Podpät'ová a nadpät'ová ochrana musí byť trojfázová. Trojfázová kontrola napätia je potrebná, aby bolo možné s istotou rozpoznať aj jednopólové poklesy napätia.

Oneskorenie vypínania podpät'ovou a nadpät'ovou ochranou musí byť krátke (okolo 100 ms), aby ani pri rýchlych zmenách napätia nedošlo ku škodám na zariadeniach ďalších odberateľov alebo na zariadení zdroja. Pri samobudení asynchrónneho generátora môže svorkové napätie behom niekoľkých periód dosiahnuť tak vysoké hodnoty, že nie je možné vylúčiť poškodenie prevádzkovaných zariadení.

Prevádzkovateľ zdroja je sám povinný zaistiť, aby spínania, kolísanie napätia, krátkodobé prerušenia, ako opätovné zapínanie, alebo iné prechodné javy v distribučnej sústave nevedli ku škodám na jeho zariadeniach.

Pre jednofázové pripojenie s výkonom do 4,6 kW nie je nutná samostatná sieťová ochrana, ak použitý striedač umožňuje nastaviť napät'ové a frekvenčné ochrany podľa požiadaviek

prevádzkovateľa distribučnej sústavy. V prípade rozšírenia existujúceho zariadenia je potrebné inštalovať samostatnú sieťovú ochranu oddelenú od striedača [1].

Aby zdroj a distribučná sústava spolupracovali spoľahlivo a bezpečne, je potrebné nastaviť ochranu zdroja tak, aby vhodne spolupracovala s ochranou distribučnej sústavy.

Pre vhodnú koordináciu ochrán je potrebné zabezpečiť:

- Pri zdrojoch pripojených do distribučnej sústavy musí prevádzkovateľ zdroja, teda výrobca elektriny, dodržať vypínacie časy poruchového prúdu tečúceho do distribučnej sústavy, aby sa dôsledky porúch v zariadeniach výrobcu prejavili v distribučnej sústave v čo najmenšom rozsahu.
- Prevádzkovateľ distribučnej sústavy zaistí, aby nastavenie ochrán vo výrobe spĺňalo vlastné vypínacie časy distribučnej sústavy.
- Na nastavení ochrán ovládajúcich vypínače alebo na nastavení automatického spínacieho zariadenia (záskoku) v ktoromkoľvek bode pripojenia k distribučnej sústave sa písomne dohodnú prevádzkovateľ DS a prevádzkovateľ zdroja v priebehu konzultácií pred pripojením. Dohodnuté hodnoty nastavenia sa nesmú meniť bez súhlasu prevádzkovateľa DS.
- Pri ochrane zdroja je potrebné zabezpečiť koordináciu s prípadnými automatikami opätovného zapínania, ktoré sú špecifikované prevádzkovateľom DS.
- Ochrany zdrojov nesmú pôsobiť pri krátkodobej asymetrii, vyvolanou likvidáciou poruchy záložnou ochranou. O veľkosti novej asymetrie napätia v sieti upovedomí prevádzkovateľ DS budúceho výrobcu elektrickej energie pri prejednávaní pripojovacích podmienok.

#### IV. NÁVRH CHRÁNENIA PRE MALÉ ZDROJE ENERGIE

Pre porovnanie návrhov ochranných funkcií boli vybrané dva druhy mini elektrární. Zdroj, ktorý vyrába energiu pomocou synchronného generátora (SG), mal inštalovaný výkon 650 kVA. Ako statický zdroj (so striedačom) bola zvolená fotovoltická elektrárňa (FVE) s výkonom 650 kW.

V oboch prípadoch bolo potrebné podľa schémy zapojenia, parametrov sústavy a elektrární vypočítať skratové pomery. Znalosť týchto pomerov je potrebná pre zadefinovanie ochranných funkcií. Maximálny a minimálny skratový prúd boli určené nasledovne:

- maximálny skratový príspevok generátora:  $I''_{k3SG}=6,2 \text{ kA}$ ,
- minimálny skratový príspevok generátora:  $I''_{k2SG}=5,4 \text{ kA}$ .

##### A. Návrh ochranných funkcií

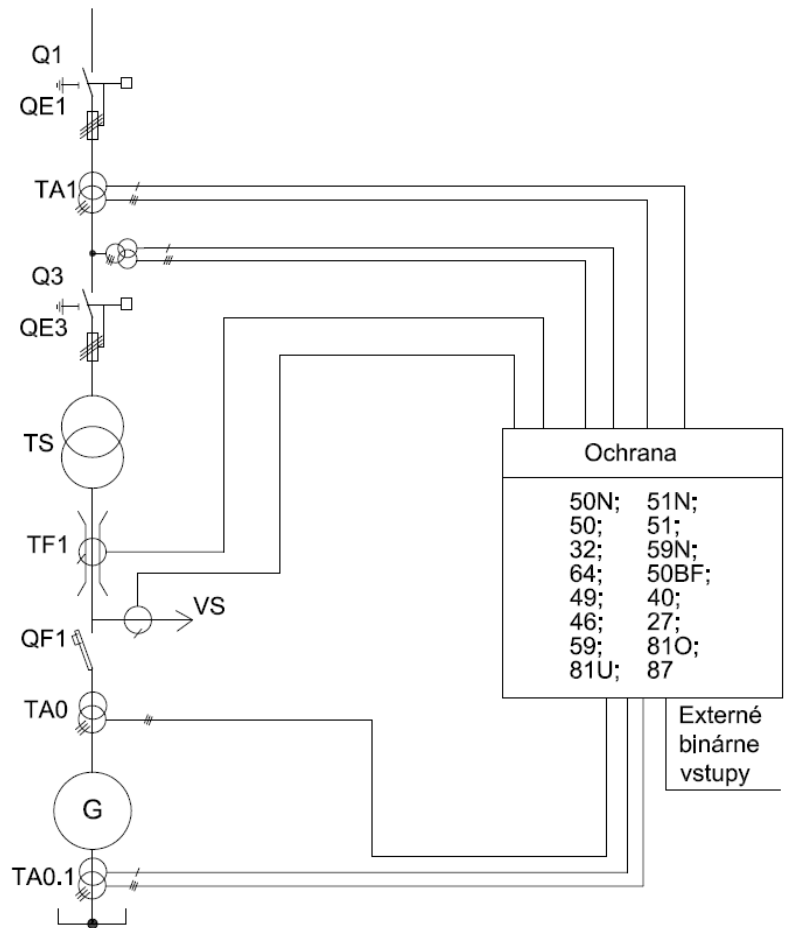
Pri sieťových ochránach je potrebné sa zamerať na kontrolu napätia a frekvencie. Hodnoty napätia aj frekvencie zdroja sa môžu hýbať len v rozmedzí určenom prevádzkovateľom distribučnej sústavy.

Pri chránení zdroja je potrebné kontrolovať viacero faktorov [3]. Sledované sú následné poruchové udalosti:

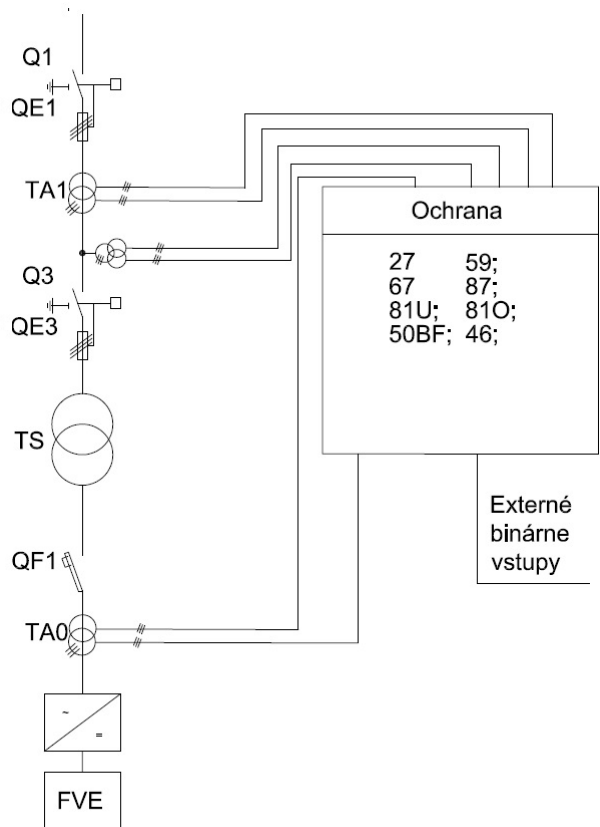
- fázové poruchy,
- zemné poruchy,
- strata budenia,
- prebudenie,
- nadpätie,
- prúdová nesymetria,
- podfrekvencia a nadfrekvencia,
- motorický chod generátora,
- poruchy vypínačov,
- systémové poruchy.

Okrem týchto základných ochranných funkcií je možné pridať ďalšie ochranné funkcie podľa potreby a finančných nárokov zdroja. Rozšírenými ochrannými funkciami môžu byť napríklad strata synchronizmu, preťaženie generátora, tepelná ochrana statora či rotora a ďalšie.

Použitie ochranné funkcie pre výrobu so SG v jednej ochrane (sieťová ochrana + ochrana generátora) sú na Obr. 1. Schéma pre chránenie FVE je uvedená na Obr. 2.



Obr. 1 Ochranné funkcie pre výrobu so SG



Obr. 2 Ochranné funkcie pri chránení FVE

V prípade energetických zdrojov pracujúcich na princípe fotovoltických článkov odpadá príspevok skratového prúdu od zdroja, nakoľko fotovoltické panely nie sú schopné dodať väčšie množstvo elektrickej energie ako je ich maximálny inštalovaný výkon. Príspevok siete je však rovnaký ako v prípade výroby so SG. Pre FVE je potrebné nastaviť sieťové ochranné funkcie tak, že v prípade poruchy v priestore medzi sieťou a zdrojom bude táto porucha zaznamenaná nadprúdovou smerovou ochrannou funkciou. Pri poruche na vývodoch z elektrárne elektrická sieť prispieva energiou smerom do elektrárne. Takáto porucha bude zistená pomocou ochrannej funkcie sledujúcej veľkosť a smer toku prúdu, ktorá po zistení poruchy rozpojí kontakty výkonového vypínača. Po rozpojení kontaktov výkonového vypínača dochádza k odpojeniu fotovoltických článkov, a tak je výkon elektrárne nulový.

## V. ZÁVER

Pre porovnanie pri návrhu ochranných funkcií boli vybrané dve mini elektrárne s rôznym princípom výroby energie. Prvá elektráreň pracuje na princípe synchronného generátora, čiže točivého stroja. Druhá je fotovoltická elektráreň, ktorá je statická a na dodávanie elektrickej energie do distribučnej siete potrebuje ešte striedač.

Pri návrhu ochrán sa určia skratové pomery v mieste pripojenia, pričom je potrebné poznať osobitne príspevok zo siete i od pripájaného zdroja.

Podľa vypočítaných hodnôt a podmienok pre pripojenie zdroja boli navrhnuté nasledovné základné ochranné funkcie s ANSI kódmi: 50N, 51N, 50, 51, 32, 59N, 64, 50BF, 49, 40, 27, 59, 81U/O. Tieto funkcie je možné doplniť o rozdielovú ochrannú funkciu (87).

Pre chránenie takejto energetickej výroby je možné použiť ochrany od viacerých výrobcov inteligentných elektronických zariadení. Ako príklad je možné uviesť:

- Siemens – 7UM621 Basic
- ABB – REG670/REG650
- Areva, resp. Schneider – MiCOM P342/P343
- SEL – 300G
- VAMP – 265, 210

V prípade ochrany FVE mnohé z ochranných funkcií použitých pri chránení synchronného stroja sú nadbytočné a nepotrebné. Bolo potrebné navrhnuť len sieťové ochranné funkcie (kontrola napätia a frekvencie) a smerovú nadprúdovú ochrannú funkciu. V tomto prípade je teda možné použiť jednoduché ochrany.

V oboch prípadoch ochrán je potrebné sledovať nie len ochranné funkcie, ale aj cenu ochrán. Je dôležité, aby cena ochrany nebola príliš vysoká vzhľadom na cenu a veľkosť samotnej elektrárne.

## LITERATÚRA

- [1] *Technické podmienky prevádzkovateľa distribučnej sústavy*, Východoslovenská distribučná, Košice, 2013.
- [2] M. Hvizdoš, "Podmienky pre pripojenie decentralizovaných zdrojov elektriny do distribučnej sústavy z hľadiska chránenia," in *Elektroenergetika*, Vol. 5, No. 2, 2012, s. 45-47, ISSN 1337-6756.
- [3] STN 333051:1992, *Ochrany elektrických strojov a rozvodných zariadení*.

# Overovanie činnosti ochrán transformátorov

<sup>1</sup>Martin VASILENKO, <sup>2</sup>Marek HVIZDOŠ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>martin.vasilenko@student.tuke.sk, <sup>2</sup>marek.hvizdos@tuke.sk

**Abstrakt** — Príspevok obsahuje rozbor porúch transformátorov a popis možností ich chránenia. Použitie nadprúdovej a diferenciálnej ochrany transformátorov je zhodnotené aj z hľadiska ekonomickej efektívnosti. Uvedený je postup nastavenia digitálnej diferenciálnej ochrany SPAD346C, ktorého súčasťou je výpočet parametrov potrebných na nastavenie vypínacej charakteristiky. Činnosť digitálnej ochrany bola následne overovaná pomocou testovacieho zariadenia CMC. Ako možná alternatíva k digitálnej ochrane SPAD346C je popísaná a otestovaná aj staršia elektromechanická rozdielová ochrana R21T.

**Kľúčové slová** — chránenie transformátorov, diferenciálna ochrana, nadprúdová ochrana, parametrizácia, testovanie

## I. ÚVOD

Porucha na transformátore je nežiaduci stav, nakoľko môže zapríčiniť nesprávnu činnosť transformátora, resp. pri väčšej poruche aj nefunkčnosť transformátora. Existuje viacero delení porúch transformátorov, no jedno z najpoužívanejších je delenie porúch na poruchy interné (vnútorné) a na poruchy externé (vonkajšie).

Interné poruchy môžeme rozdeliť do dvoch skupín:

- poruchy s okamžitou detekciou,
- poruchy vznikajúce postupne.

Externé poruchy môžeme rozdeliť takisto do dvoch skupín:

- preťaženie,
- vonkajší skrat.

Medzi poruchy s možnosťou okamžitej detekcie možno zaradiť skraty na svorkách, skraty na vinutí transformátora, skraty medzi vinutiami, zemné poruchy a závitové skraty terciárnych vinutí. Medzi poruchy vznikajúce postupne možno zaradiť poruchy chladenia, chybnú reguláciu, nevhodnú, resp. nekvalitnú izoláciu plechov [1].

Nakoľko transformátor z hľadiska nákladov patrí medzi najdrahšie elektrické zariadenia, je potrebné transformátor chrániť, aby sa v prípade poruchy nepoškodil, ale bol včas vypnutý. Na tento účel nám slúžia elektrické ochrany. Princíp elektrickej ochrany spočíva v detekcii buď veličiny neelektrickej, alebo veličiny elektrickej.

Typickým predstaviteľom ochrán, ktoré reagujú na neelektrické veličiny, je tzv. Buchholzovo relé. Reaguje na plyny, ktoré pri poruchách vznikajú v určitých miestach transformátora [2].

Princíp transformátorovej ochrany Sergi je v oddelení plynov, odvedení plynov mimo oblasť transformátora a zabránení vývinu výbušných plynov pomocou injektáže dusíka [3].

Princíp nádobovej ochrany spočíva v meraní prúdu tečúceho cez uzemnenie nádoby transformátora.

Úlohou ochrany pred zvýšeným magnetickým tokom (ochrana V/Hz, resp. U/f) je chrániť jadro transformátora [1].

Princíp nadprúdovej ochrany je v pôsobení v nastavenom čase vtedy, ak je hodnota niektorého z nameraných prúdov väčšia ako hodnota nastavená na ochrane.

Princíp rozdielovej (diferenciálnej) ochrany je založený na predpoklade, že cez chránený objekt prevádzkové prúdy len prechádzajú. Za normálneho stavu je teda súčet prúdov vstupujúcich a prúdov vystupujúcich z objektu je rovný 0. Ak súčet daných prúdov nie je nulový, predstavuje to existenciu poruchového miesta, ktorým uniká rozdielový prúd [4].



## II. POROVNANIE POUŽITIA OCHRÁN TRANSFORMÁTOROV Z EKONOMICKÉHO HĽADISKA

Norma STN 333051 pri sieťových transformátoroch s výkonom od 1,7 MVA do 25 MVA nedefinuje jednoznačne použitie nadprúdovej alebo rozdielovej ochrany transformátora [5]. Z tohto dôvodu je vhodné porovnať na základe ekonomických kritérií, použitie tej či onej ochrany. Do presného ekonomického prepočtu je potrebné zahrnúť veľa premenných, ktorých hodnota môže byť nejednoznačná a ťažko vyčísliteľná. Takisto je potrebné porovnávať porovnateľné s porovnateľným a výpočet je potrebné zjednodušiť vopred stanovením rozsahu hodnôt premenných veličín.

Na hodnotenie investičných variantov s krátkym časom výstavby a rovnakou životnosťou možno použiť statickú nákladovú metódu založenú na porovnávaní prevádzkových a jednorazových, čiže investičných nákladov [6]. Porovnávať budeme pomocou absolútnych hodnôt ich celkových výrobných nákladov  $N_v$  podľa vzťahu:

$$N_v = N_i + \sum_{t=1}^{Tz} N_{pr} \quad (1)$$

kde  $N_i$  sú celkové investičné náklady,  
 $N_{pr}$  sú prevádzkové náklady.

Prepočet bol vykonaný pre 2 typy transformátorov, a to pre transformátor s menovitým výkonom  $S = 2$  MVA a transformátor s menovitým výkonom  $S = 20$  MVA. Pre oba transformátory sa prepočítali náklady v prípade, ak nenastane žiadna porucha, a pre prípad najhoršej možnej poruchy – medzivinuťového skratu, kedy bude potrebné poškodené časti transformátora vymeniť. Pre poruchu bolo určené, že nastane v piatom roku prevádzky. Investičné náklady vo výpočte pozostávajú z ceny za terminál s ochranou a z ceny za transformátor. Prevádzkové náklady pozostávajú z nákladov za revízie a z ceny komponentov transformátora, ktoré by bolo potrebné vymeniť v prípade poškodenia pri poruche. Ceny transformátorov a jednotlivých komponentov boli získané od výrobcu.

Tab. 1 Porovnanie celkových výrobných nákladov  $N_v$

	Terminál s nadprúdovou ochranou	Terminál s diferenciálnou ochranou	Porovnanie
<b>Transformátor S = 2 MVA</b>			
Normálna prevádzka	$N_{v1} = 20600$ EUR	$N_{v2} = 22200$ EUR	$N_{v1} < N_{v2}$
Najhoršia možná porucha	$N_{v1} = 35400$ EUR	$N_{v2} = 32360$ EUR	$N_{v1} > N_{v2}$
<b>Transformátor S = 20 MVA</b>			
Najhoršia možná porucha	$N_{v1} = 683400$ EUR	$N_{v2} = 586400$ EUR	$N_{v1} > N_{v2}$

Na základe vyššie uvedených prepočtov je možné vysloviť tvrdenie, že v prípade, ak počas prevádzky nenastane závažná porucha chránenom transformátore, je použitie variantu s terminálom nadprúdovej ochrany výhodnejšie. V prípade, ak počas prevádzky závažná porucha nastane, variant s použitím terminálu s diferenciálnou ochranou sa javí ekonomicky výhodnejší, z dôvodu, že škody (náklady) budú menšie, ako v prípade variantu s nadprúdovou ochranou. Rozdiel v nákladoch na škodu je ešte viac viditeľný v prípade drahších transformátorov.

Vo výpočte bolo definované, kedy nastane porucha, no v praxi nie je možné vydedukovať, kedy daná porucha nastane a či vôbec nastane. Preto je vhodné vychádzať zo štatistických údajov a tie zahrnúť do výpočtu.

Poruchovosť v dôsledku vyskytnutia skratu medzi vinutiami je v prípade distribučných transformátoroch je 0,68 %, resp. 0,7 % po zaokrúhlení (merná poruchovosť je 0,7 poruchy na 100 ks prevádzkovaných transformátorov). V prípade zakomponovania tohto štatistického údaju do celkových nákladov dostaneme výsledky v Tab. 2.

Tab. 2 Porovnanie celkových výrobných nákladov  $N_v$

	Terminál s nadprúdovou ochranou	Terminál s diferenciálnou ochranou	Porovnanie
<b>Transformátor S = 2 MVA</b>			
Najhoršia možná porucha	$N_{v1} = 19512$ EUR	$N_{v2} = 21800$ EUR	$N_{v1} < N_{v2}$
<b>Transformátor S = 20 MVA</b>			
Najhoršia možná porucha	$N_{v1} = 345780$ EUR	$N_{v2} = 346690$ EUR	$N_{v1} < N_{v2}$

V prípade zakomponovania štatistických údajov do výpočtu je možné vysloviť tvrdenie, že je výhodnejšie použiť terminál s nadprúdovou ochranou, no s narastajúcim výkonom transformátora, narastajúcou cenou transformátora, ako aj s narastajúcou poruchovosťou by mohli byť náklady menšie v prospech terminálu s diferenciálnou ochranou.

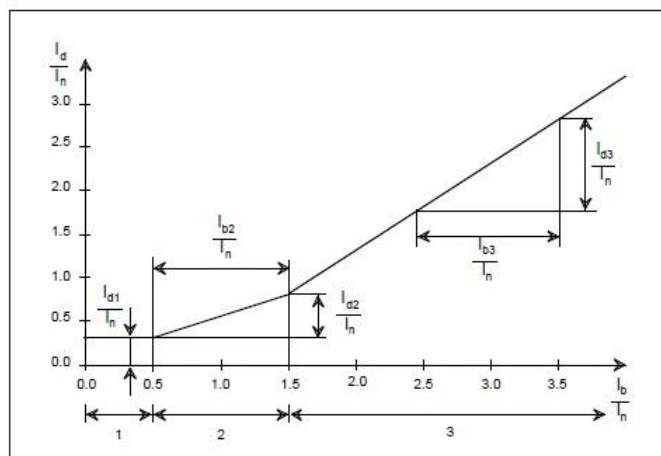
### III. PARAMETRIZÁCIA A OVERENIE ČINNOSTI OCHRÁN TRANSFORMÁTORA

Parametrizácia a následne aj overenie bolo realizované na rozdielovej digitálnej ochrane SPAD346C a na elektromechanickej ochrane R21T. V prípade digitálnej ochrany je ochrana parametrizovaná pre dva prípady reálnych transformátorov z prevádzky (označené ako Tr1 a Tr2).

#### A. Digitálna rozdielová ochrana SPAD346C

Diferenciálna ochrana SPAD346C je ochrana, ktorá je určená na chránenie dvojvinuťových transformátorov, blokov generátor – transformátor a generátorov. Pri transformátoroch sa používa na chránenie pri medzizávitových poruchách, pri zemných poruchách a skratoch medzi vinutiami. Pri generátoroch sa používa na chránenie pri závitových a medzifázových skratoch. Ochrana môže byť použitá aj pri chránení trojvinuťových transformátorov v prípade, ak skratový výkon je zo 75 % dodávaný z rovnakého smeru.

Na Obr. 1 je znázornená vypínacia charakteristika tejto ochrany. Z danej charakteristiky je možné vidieť, ktorý parameter ako ovplyvňuje výsledné nastavenie charakteristiky [7].



Obr. 1 Vypínacia charakteristika diferenciálnej ochrany SPAD346C

Pokiaľ menovité hodnoty prúdu výkonového transformátora na primárnej a sekundárnej strane nie sú rovnaké ako menovité prúdy prístrojových transformátorov prúdu (PTP) na primárnej a sekundárnej strane, potom sa použijú parametre nastavenia  $I_1/I_n$  a  $I_2/I_n$  nazývané aj korekcia transformačných prevodov. Korekciu transformačných pomerov si vypočítame ako podiel prúdu daného vinutia k hodnote primárneho prúdu PTP.

Hodnota korekcie transformačného prúdu pre primárny prúd Tr1:

$$\frac{I_1}{I_n} = \frac{209,95}{250} = 0,84 \quad (2)$$

Hodnota korekcie transformačného prúdu pre sekundárny prúd Tr1:

$$\frac{I_2}{I_n} = \frac{1004,1}{1250} = 0,80 \quad (3)$$

Prúdový pomer  $S$  zahŕňa presnosť prístrojových transformátorov, odbočky, chyby prispôsobenia a zálohu. Pre Tr1 je:

$$S = (5 + 5) + (16) + (22) + (5) = 35 \% \quad (4)$$

Parameter  $P/I_n$  pre Tr1:

$$\frac{P}{I_n} = \frac{I_b}{I_n} \cdot S = 0,5 \cdot 35 + 20 = 38 \% \quad (5)$$

Nastavenie bodu druhého zlomu vypínacej charakteristiky:

$$I_{2pt}/I_n = 2 \quad (6)$$

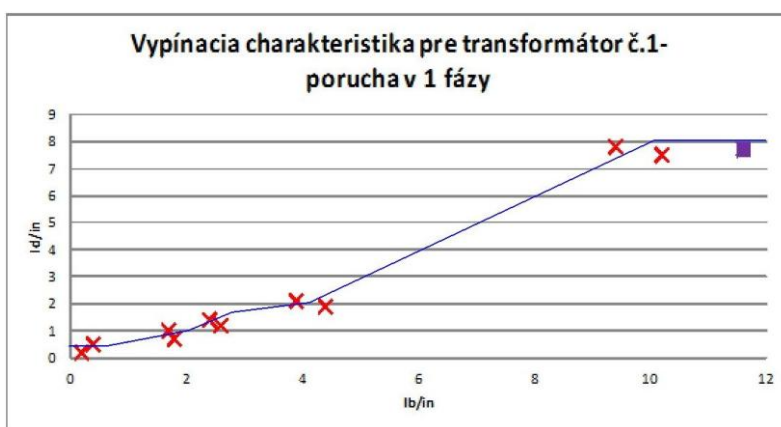
Tak isto sa postupuje aj pri výpočte parametrov pre Tr2. Vypočítané parametre sú uvedené Tab. 3.

Tab. 3 Vypočítané parametre pre transformátory Tr1 a Tr2

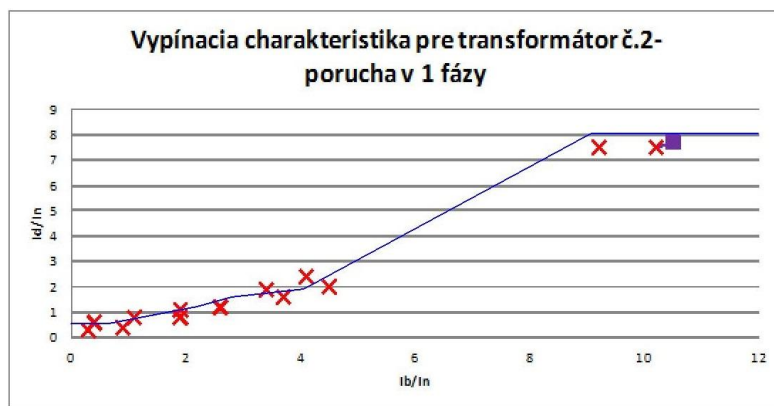
Parametre	Transformátor Tr1	Transformátor Tr2
$I_1/I_n$ [-]	0,84	0,88
$I_2/I_n$ [-]	0,80	0,96
$S$ [%]	35	34
$P/I_n$ [%]	38	42
$I_{2TP}/I_n$ [-]	2	2

Vypínacia charakteristika bola testovaná pre body zvolené na základe poznatkov z praxe, kde sa charakteristika testuje v zlomoch vypínacej charakteristiky.

Skúšobným zariadením CMC bolo otestované pôsobenie ochrany na zmenu prúdu v prvej fáze (Obr. 2 a Obr. 3). Ochrana zareagovala správne vo všetkých prípadoch, pričom posledný testovaný bod bol mimo prúdového rozsahu testovacieho zariadenia.



Obr. 2 Vypínacia charakteristika ochrany SPAD346C pre transformátor Tr1

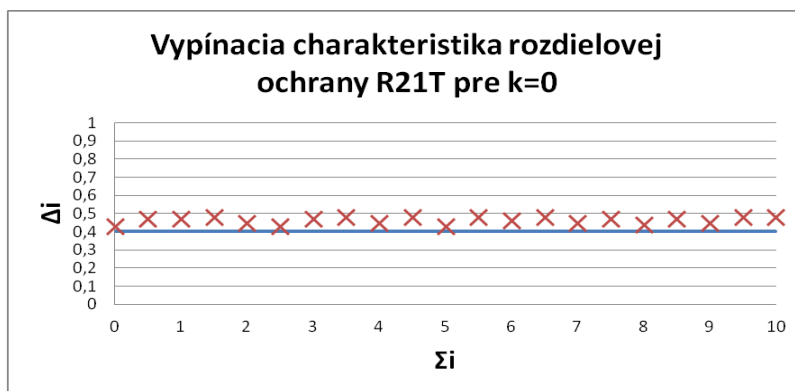


Obr. 3 Vypínacia charakteristika ochrany SPAD346C pre transformátor Tr2

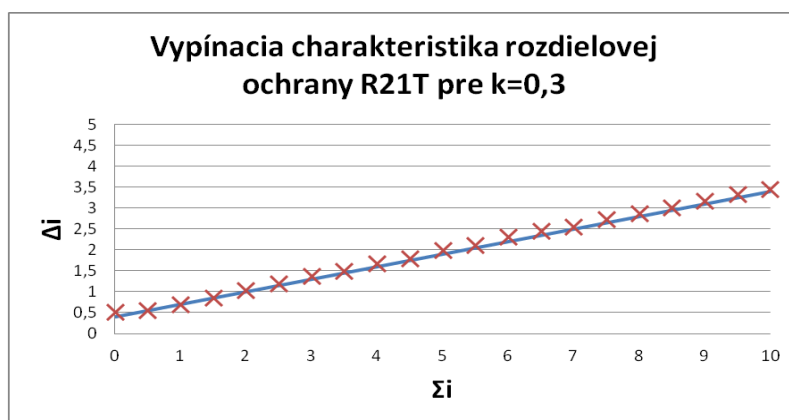
### B. Elektromechanická rozdielová ochrana R21T

Ochrana typu R21T je elektromechanická ochrana určená na detekciu vnútorných porúch transformátorov. V súčasnosti sa už v praxi nepoužíva, bola nahradená elektromechanickou ochranou R30. Tá sa v súčasnosti ešte používa, ale je nahradzovaná digitálnymi rozdielovými ochranami. Pre ochranu R21T je dôležité, aby prúd pred chráneným objektom  $I_1$ , prúd za chráneným objektom  $I_2$  a ich rozdiel v ustálenom stave  $\Delta i$  boli vo fáze, resp. aby hodinový uhol bol rovný 0. V prípade, ak je hodinový uhol nie je rovný 0 (kvôli zapojeniu vinutí transformátora), používajú sa medzitransformátory prúdu [8].

Na Obr. 4 je znázornená vypínacia charakteristika s nastaveným znižovacím faktorom  $k$  na hodnotu 0. Ako je vidieť z charakteristiky, ochrana pôsobila v hodnote o niečo vyššej ako bola nastavená hodnota na ochrane. Na Obr. 5 je znázornená vypínacia charakteristika so znižovacím faktorom  $k$  nastaveným na hodnotu 0,3.



Obr. 4 Vypínacia charakteristika rozdielovej ochrany R21T pre  $k = 0$



Obr. 5 Vypínacia charakteristika rozdielovej ochrany R21T pre  $k = 0,3$

#### IV. ZÁVER

Pre chránenie transformátorov je rozhodujúca norma STN 333051. Tá podľa výkonu transformátora určuje, či je nutné použiť na jeho chránenie rozdielovú ochranu, alebo postačuje nadprúdová ochrana. Pre výkony od 1,7 MVA do 25 MVA je rozdielová ochrana len odporúčaná. Kritériom pre použitie je teda ekonomické hľadisko. Výhodnosť použitia nadprúdovej a diferenciálnej ochrany bola posudzovaná pre stav bez poruchy a pre prípad najhoršej možnej poruchy, čiže medzivinuťového skratu. Výpočet bol realizovaný pre dva transformátory, a to pre transformátor s menovitým výkonom 2 MVA a transformátor s výkonom 20 MVA. Z ekonomického prepočtu vyplýva, že v prípade drahších transformátorov má zmysel použiť terminál s diferenciálnou ochranou. V prípade zakomponovania štatistického údaja o možnom výskyte poruchy transformátora (do 20 MVA) je možné vysloviť tvrdenie, že použitie terminálu s nadprúdovou ochranou je výhodnejšie.

Cieľom práce bolo zároveň parametrizovať a overiť činnosť ochrán transformátorov. Pre účely merania bola zvolená digitálna rozdielová ochrana SPAD346C od firmy ABB. Činnosť tejto digitálnej rozdielovej ochrany bola porovnávaná so staršou elektromechanickou rozdielovou ochranou R21T.

Ochrana SPAD346C bola parametrizovaná pre dva konkrétne transformátory používané v prevádzke. Na základe manuálu ochrany boli vypočítané a nastavené hodnoty, ktorými sa vypínacia charakteristika ochrany nastavuje. Následne bola vypínacia charakteristika ochrany otestovaná pomocou testovacieho zariadenia CMC256plus. Ochrana v jednotlivých bodoch charakteristiky pôsobila správne v čase na to určenom, čiže do 50 ms.

Následne ako možná alternatíva k ochrane SPAD 346C bola parametrizovaná ochrana R21T. Hodnoty sa nastavili ručne na ochrane a ochrana sa testovala pre dva typy charakteristík. Ochrana bola otestovaná skúšobným zariadením CMC156. Ochrana v jednotlivých bodoch vypínacej charakteristiky pôsobila pri hodnote o niečo vyššej ako nastavená hodnota, čo mohlo byť spôsobené celkovou zastaranosťou ochrany.

Overenie činnosti rozdielových ochrán ukázalo, že digitálna rozdielová ochrana SPAD346C pôsobí rýchlejšie, presnejšie a spoľahlivejšie ako elektromechanická ochrana R21T. Z daného vyplýva, že digitálne rozdielové ochrany sú pre prax vhodnejšie, pričom ponúkajú možnosť podrobnejšej parametrizácie ochrany pre danú vypínaciu charakteristiku.

LITERATÚRA

- [1] F. Janiček, V. Chladný, A. Beláň, Ž. Eleschová, *Digitálne ochrany v elektrizačnej sústave*. STU Bratislava, 2004, 360 s., ISBN 80-227-2135-2.
- [2] *Buchholz Relay – Construction & Operation of Buchholz Relay*. [online] [cit. 2014-1-20] dostupné na internete: <<http://www.electrical4us.com/buchholz-relay-definition-construction-working-principle-of-buchholz-relay/>>
- [3] *Transformer Protector*. SERGI.
- [4] J. Tlustý et al., *Monitorování, řízení a chránění elektrizačních soustav*. ČVUT Praha, 2011, 256 s., ISBN 978-80-01-04940-2
- [5] STN 333051:1992, *Ochrany elektrických strojov a rozvodných zariadení*.
- [6] A. Mészáros, *Ekonomika elektroenergetiky*. TU Košice, 2013, 206 s., ISBN 978-80-553-1512-6.
- [7] *SPAD346C – Stabilized Differential Relay*. ABB.
- [8] M. Hvizdoš, *Elektrické ochrany v elektrizačnej sústave – Návod na cvičenia*. TU Košice, 2012, 70 s. ISBN 978-80-553-1193-7.

# Chránenie trakčných napájacích staníc

<sup>1</sup>Rastislav TOMÁŠKO, <sup>2</sup>Marek HVIZDOŠ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>rastislav.tomasko@student.tuke.sk, <sup>2</sup>marek.hvizdos@tuke.sk

**Abstrakt** — Príspevok sa zaoberá chránením trakčných napájacích staníc. Tieto sú súčasťou trakčnej sústavy. Uvedené je rozdelenie, spôsoby napájania a popis trakčných napájacích staníc. Ďalej sa zaoberá skratmi, príčinami ich vzniku, rozdelením a podmienkami pre výpočet skratových prúdov. Rieši výpočet skratových prúdov v dôležitých miestach stanice. Pre určené oblasti chránenia je uvedený návrh nastavenia ochrán konkrétnej napájacej stanice.

**Kľúčové slová** — trakčná napájacia stanica, skrat, ochrana

## I. ÚVOD

Trakčné napájacie stanice patria medzi dôležité zariadenia, ktoré zabezpečujú potrebnú elektrickú energiu pre napájanie trakčného vedenia. Spájajú rozvodnú energetickú sieť s trakčnou sieťou a zaisťujú dodávku energie v potrebnej forme pre trakčné účely.

Z hľadiska svojej funkcie zahrňujú trakčné napájacie stanice všetky elektrické zariadenia od vstupu prívodného vedenia, až k miestu výstupu napájacieho vedenia, spájajúceho prípojnice trakčnej napájacej stanice s trolejovým a koľajnicovým vedením.

## II. TRAKČNÁ SÚSTAVA

Elektrické trakčné zariadenia sú typickým elektrickým silovým zariadením v doprave. V prípade závislej trakcie závisí funkcia koľajového vozidla od vonkajšieho prívodu energie. Vozidlo so sebou nevozi energetický zdroj, nemusí byť odstavované z prevádzky pre dopĺňanie pohonnými hmotami. Elektrická trakcia efektívne využíva pre pohon energiu, ktorá je vyrobená z menej hodnotných prírodných palív, vodných síl a jadrového paliva vo veľkých elektrárňach, ktorých prevádzka je energeticky účinnejšia, hospodárnejšia a bezpečnejšia. Závislá trakcia je najracionálnejším druhom pozemnej dopravy.

Elektrická energia sa prenáša na elektrické hnacie vozidlá pomocou pevných elektrických trakčných zariadení (PETZ). Z nadradenej trojfázovej siete VN alebo VVN sa energia privádza do trakčných napájacích staníc (TNS), ktoré sú umiestnené v blízkosti trate.

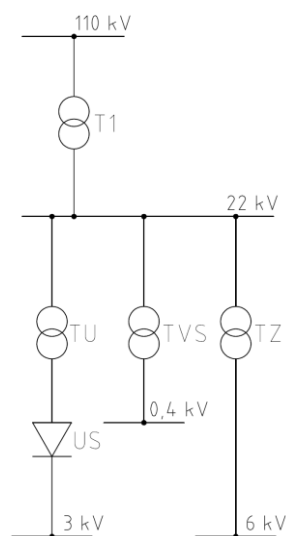
### A. Trakčné napájacie stanice

Trakčné napájacie stanice (TNS) sú elektrické stanice, ktoré slúžia na úpravu napájacieho napätia trojfázovej napäťovej sústavy VVN alebo VN na trakčnú napájajúcu napäťovú sústavu vhodnú pre napájanie elektrických hnacích dráhových vozidiel.

Podľa druhu trakčnej napájacej napäťovej sústavy sa TNS delia na:

- trakčné meniarne - napájajú elektrifikované trate železníc s jednosmernou napäťovou sústavou 3 kV,
- trakčné transformovne – napájajú elektrifikované trate železníc so striedavou jednofázovou napäťovou sústavou 25 kV, 50 Hz.

V meniarňach sa trojfázové striedavé napätie transformuje a usmerňuje na napätie jednosmernej trakčnej napäťovej sústavy 3 kV. V transformovniach sa združené vstupné napätie znižuje na príslušnú hodnotu jednofázovej trakčnej napäťovej sústavy 25 kV, 50 Hz. Menovité napätie na prípojniciach rozvodne v trakčnej napájacej stanici musí byť vždy vyššie ako menovité napätie trakčnej napäťovej sústavy. Menovité trakčné napätie vyjadruje hodnotu napätia na zberači hnacieho vozidla, ku ktorému sa viažu charakteristiky a menovité technické údaje hnacích vozidiel a ich elektrických zariadení. Z napájacích staníc energia postupuje ku zberačom hnacích vozidiel obvodom, ktorý nazývame trakčné vedenie.



#### Hlavné časti TNS:

- rozvodňa 110 kV,
- rozvodňa 22 kV,
- rozvodňa 6 kV,
- rozvodňa napájacieho napätia 3 kV,
- usmerňovacia skupina,
- rozvádzač vlastnej spotreby 0,4 kV.

Obr. 1 Hlavné časti trakčných napájacích staníc

Z hľadiska smeru toku sú trakčné napájacie stanice pripojené vstupnou časťou na rozvodnú sieť distribučnej spoločnosti a výstupnou časťou na trakčné vedenie. Trakčné vedenie je pevné elektrické zariadenie, ktoré slúži na privedenie elektrickej energie do hnacieho dráhového vozidla z trakčnej napájacej stanice. Trakčné vedenie, ktorým sa privádza elektrická energia z trakčnej napájacej stanice k elektrickým hnacím dráhovým vozidlám je rozdelené na jednotlivé navzájom elektricky izolované úseky, tzv. napájacie úseky. Dĺžka napájacieho úseku závisí od druhu použitej trakčnej napájacej napät'ovej sústavy a od spôsobu napájania.

Pre obe trakčné napät'ové sústavy sa musia dimenzovať pevné elektrické trakčné zariadenia tak, aby napájacie napätie trakčného vedenia pri prevádzke kolísalo len v dovolených medziach. Menovité napätie trakčných meniarň sa meria na prípojnicach rozvodne jednosmerného napätia pri menovitom zaťažení meniarne a pri menovitom vstupnom napätí na primárnom vinutí usmerňovačových transformátorov. Z toho dôvodu, že vstupné napätie vvn alebo vn kolíše, trakčné transformátory v trakčnej napájacej stanici majú regulovateľné odbočky, ktoré obmedzujú vplyv kolísania napájania a trakčné vedenie je napájané napätím v predpísanej tolerancii [1], [2], [3].

### III. PORUCHY V TRAKČNÝCH NAPÁJACÍCH STANICIACH

Pri prevádzke trakčných napájacích staníc je potrebné počítať nielen s ustáleným prevádzkovým stavom, ale aj s prechodnými stavmi, medzi ktoré patria poruchy, ako sú:

- skrat,
- prepätie,
- zemné spojenie.

Práve takéto poruchy nám môžu poškodiť elektrické zariadenia nachádzajúce sa v trakčnej napájacej stanici.

V mieste skratu dochádza k poklesu napätia, pričom do miesta poruchy tečú skratové prúdy zo všetkých zdrojov sústavy, v závislosti od ich výkonu a elektrickej vzdialenosti od miesta poruchy. Úbytky napätia vzrastajú od miesta skratu smerom ku zdrojom, a preto pokles napätia sa môže prejaviť prakticky v celej elektrizačnej sústave.

#### A. Rozdelenie skratov

Trojfázový skratový prúd je spravidla najväčší. Na vonkajších vedeniach sa vyskytuje zriedkavo, ale napriek tomu, je potrebné sa ním zaoberať, keďže elektrické zariadenia sa dimenzujú na najvyšší skratový prúd. Na rozdiel od vonkajších vedení sa u káblových vedení vyskytujú skoro vo všetkých prípadoch trojfázové skraty, pričom oblúk poruší izoláciu všetkých troch fáz. Elektrické zariadenia musia taktiež odolať aj iným druhom skratov, ktoré spôsobujú najväčšie namáhanie. Môže to byť dvojfázový skrat na vedeniach, ktoré spájajú jednofázové transformátory s trojfázovými prípojnicami. Najvyšší stupeň rušenia zariadení spôsobuje jednofázový skrat.

Nepriaznivé účinky skratových prúdov sú o to vážnejšie, že sa často prejavujú súčasne. Najčastejšie sledované sú mechanické a tepelné účinky, ktoré môžeme obmedziť správnym

dimenzovaním elektrických zariadení. Pokles napätia v dôsledku skratu má vplyv na svetelné zdroje, na tepelné spotrebiče, ale hlavne ohrozuje prevádzku motorov. Náhla zmena zaťaženia v dôsledku skratu môže vážne ohroziť stabilitu prenosu elektrickej energie.

S cieľom minimalizácie dôsledkov takýchto porúch musí byť skrat odpojený čo najrýchlejšie od zvyšku elektrizačnej sústavy. Túto funkciu v rozvodoch NN plnia ističe a poistky, v sieťach VN poistky a ochrany a v sieťach VVN a ZVN len ochrany. Po identifikácii poruchy dáva ochrana svojim koncovým členom povel na vypínanie cievku výkonového vypínača [4].

#### IV. NÁVRH NASTAVENIA OCHRÁN TRAKČNEJ NAPÁJACEJ STANICE

V poslednej dobe sa kladú stále väčšie nároky na spoľahlivosť pre všetky elektrické zariadenia. Ich počet sa zvyšuje a taktiež investičné a prevádzkové náklady týchto zariadení narastajú. Každé elektrické zariadenie je chránené ochranou, ktorá musí rozoznať, či ide o normálny alebo poruchový stav.

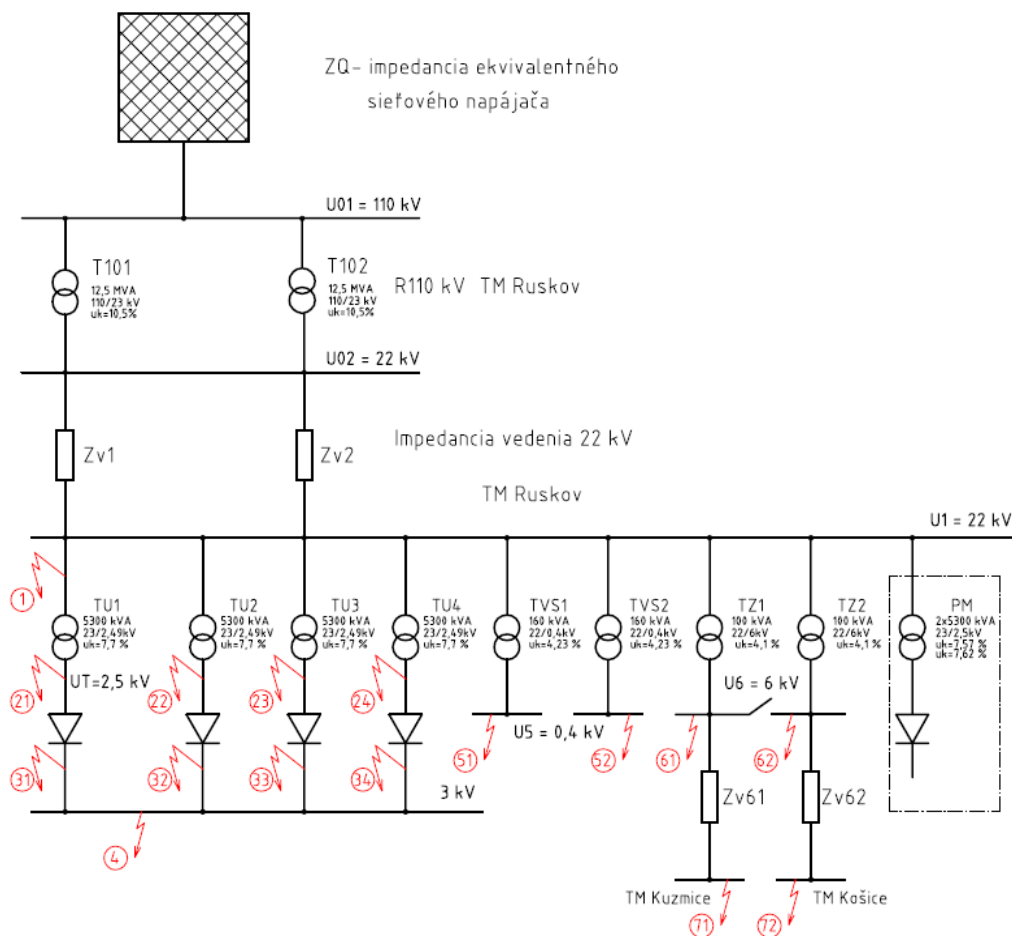
K správne nastaveniu rozbehových prúdov na ochrane potrebujeme poznať prúdové pomery, predovšetkým:

- hodnoty skratových prúdov,
- veľkosti dovolených prúdov pri preťažení,
- menovitý prúd.

##### A. Určenie skratových prúdov

Veľkosti skratových prúdov, pri tej istej konfigurácii siete, môžu byť v rôznych časoch prevádzky rôzne veľké. Veľkosť skratových prúdov závisí od výkonu zdrojov, ktoré v danom okamihu pracujú do sústavy. Z toho dôvodu potrebujeme poznať maximálny a minimálny skratový prúd, ktorý sa môže vyskytnúť v mieste zabudovania ochrany. Veľkosť týchto prúdov zisťujeme výpočtom podľa STN EN 60909.

V našom prípade boli skratové pomery určené podľa schémy na Obr. 2.



Obr. 2 Náhradná schéma pre výpočet skratových prúdov



Pri výpočte maximálnych skratových prúdov je potrebné vziať do úvahy najnepriaznivejšie pomery v chránenej sieti, teda musíme predpokladať chod všetkých transformátorov, vedení, najnepriaznivejšie miesto skratu a pod. Naopak pri výpočte minimálnych skratových prúdov je potrebné uvažovať so vznikom skratu v čase, keď pracuje do siete minimálny počet zdrojov, v čase najmenšieho zaťaženia, napr. noc, deň pracovného pokoja, a pod.

K určeniu maximálnych a minimálnych hodnôt skratových prúdov je potrebné poznať prevádzkovú schému siete. Nastavenie ochrán sa robí pre danú prevádzkovú schému siete. Toto nastavenie ochrán musí zabezpečiť selektívne vypnutie všetkých typov porúch.

### B. Parametrizácia ochrán

Hodnoty, na ktoré sa navrhujú nastavenia ochranných funkcií, musia v príslušných oblastiach chránenia spĺňať tieto podmienky:

- zaistiť vypnutie minimálneho skratového prúdu,
- zabrániť výkonovému preťaženiu chránených častí technológie,
- zabrániť nežiaducemu vypínaniu transformátorov pôsobením prúdového nárazu pri zapínaní, pričom je treba očakávať až 10 násobné prekročenie hodnoty menovitého prúdu, trvajúce až niekoľko sekúnd a použiť ochrannú funkciu spočívajúcu v detekcii 2. harmonickej v zapínacom prúde transformátora,
- zaistiť čo najvyššiu stabilitu energetického napájania dostatočným odstupom najvyšších prevádzkových prúdov od nastavených vypínacích prúdov,
- zaistiť nevyhnutnú úroveň selektivity ochranných funkcií medzi miestom skratu a zdrojom. Plnením tejto podmienky sa má súčasne dosiahnuť rýchlejší nábeh nadprúdových ochranných funkcií a rýchlejšie vypínanie výkonových vypínačov bližších k miestam skratu.

Návrh nastavenia nadprúdových ochrán pre definované oblasti chránenia v primárnych hodnotách je uvedený v Tab. 1.

Tab. 1 Návrh nastavenia ochranných funkcií

Názov chránenej zóny	Menovitý prúd chráneného zariadenia [A]	Rozhodujúci skratový prúd $I''_{k3min}$ [kA]	Funkcia 3I2f	Nastavenie v primárnych hodnotách					
				I>	t>	I>>	t>>	I>>>	t>>>
				[A]	[s]	[A]	[s]	[A]	[s]
prívod 22 kV P1-2	314	1,846	nie	314	40	408	1,5	-	-
vývod 22 kV TU1-4	133	0,946 0,600	áno	320	6	400	0,2	800	0,05
prívod 3 kV U1-4	1500	12,342	x	3000	40	3600	0,3	4800	0,07
vývod 22 kV PM	266	1,846	áno	400	1	850	0,08	-	-
vývod 22 kV VS1-2	4,2	0,095	nie	4,6	60 sig.	6,3	0,5	16,8	0,1
vývod 22 kV TZ1-2	2,62	0,062	nie	4,2	4	11	2,5	35	0,05
prívod 6 kV QZ11-12	9,62	0,228	nie	13,5	3	35	2	-	-
prívod 6 kV QZT1-2	9,62	0,107	nie	12,5	1,8	30	0,6	-	-

## V. ZÁVER

Rovnako ako dochádza k modernizácii trakčných napájacích staníc so zastaranou technológiou, tak sa postupne nahrádzajú staré elektromechanické ochrany digitálnymi ochranami, ktoré poskytujú vyššiu rýchlosť pôsobenia, presnosť, selektivitu, komplexnejšie spracovanie vstupných veličín a zvýšenie užívateľského komfortu.

Pre zaistenie bezpečnej prevádzky trakčných napájacích staníc je úloha ochrán nezastupiteľná. Z toho dôvodu sa musí ochranám venovať veľká starostlivosť a zaistenie ich trvalého bezporuchového chodu. Pre každú trakčnú napájaciu stanicu, musí byť vypracovaný plán nastavenia ochrán a harmonogram pravidelných skúšok, revízií a údržby.

LITERATÚRA

- [1] J. Trávníček, M. Foglar, F. Hudec, *Energetika v doprave*. NDaS Praha, 1988, 420 s.
- [2] G. Lanáková, V. Oslovič, *Pevné elektrické trakčné zariadenia*. EDIS Žilina, 2006, 198 s. ISBN 80-8070-507-0.
- [3] G. Lanáková, D. Šindler, *Napájanie elektrických dráh*. ALFA Bratislava, 1989, 206 s.
- [4] F. Janiček, V. Chladný, A. Beláň, Ž. Eleschová, *Digitálne ochrany v elektrizačnej sústave*. STU Bratislava, 2004, 360 s., ISBN 80-227-2135-2.

# Overovanie činnosti ochrán proti zemným spojeniam

<sup>1</sup>Lukáš KOLARČIK, <sup>2</sup>Marek HVIZDOŠ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>lukas.kolarcik@student.tuke.sk, <sup>2</sup>marek.hvizdos@tuke.sk

**Abstrakt** — Príspevok sa zaoberá ochranami proti zemným spojeniam, predovšetkým ich parametrizáciou a testovaním. Rozoberá problematiku zemných porúch a uvádza princípy na ich detekciu. Činnosť zemných ochrán bola overovaná pomocou testovacieho zariadenia CMC.

**KLúčové slová** — zemná ochrana, ochranný terminál, parametrizácia ochrán, testovanie ochrán

## I. ÚVOD

Účel a použitie zemných ochrán je rôzny. Zemné ochrany sa používajú na chránenie vývodov, vedení a generátorov. Ich použitie závisí od princípu, na akom ochrana pracuje. Vzhľadom na konštrukčné vyhotovenie, tieto ochrany rozdelujeme do dvoch kategórií:

- elektromechanické zemné ochrany
- digitálne zemné ochrany

Algoritmus klasických elektromechanických zemných ochrán je založený na meraní prúdu  $3I_0$  a napätia  $3U_0$ . Ochrany z týchto veličín zisťujú tok nulovej zložky výkonu na základe ich veľkosti a fázového uhla, ktorý vzájomne zvierajú. Elektromechanický princíp neumožňuje výpočet absolútnej hodnoty nulovej zložky výkonu, preto sú tieto ochrany vyhotovené ako smerované. Môžu byť wattové alebo jalové [1].

Princíp digitálnych ochrán je taktiež založený na meraní prúdu  $3I_0$  a napätia  $3U_0$ . Softvér týchto ochrán obohacuje možnosti ich použitia, keďže z týchto veličín dokážu zistiť okrem fázového uhla aj absolútnu hodnotu nulovej zložky výkonu. Ich výhodou je taktiež jemné nastavenie citlivosti pôsobenia [2].

Tab. 1 Ochranné funkcie zemných ochrán

Typ ochrannej funkcie	Merané veličiny	Na vyhodnotenie zemnej poruchy je potrebné
citlivá zemná smerová ochrana	$3I_0, 3U_0, \varphi_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>3U_0</math></li><li>▪ prekročenie nastavenej hodnoty <math>3I_0</math></li><li>▪ uhol <math>\varphi_0</math> medzi prúdom a napätím je v nastavenom rozmedzí</li></ul>
zemná smerová ochrana	$3I_0, 3U_0, \varphi_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>3I_0</math></li><li>▪ uhol <math>\varphi_0</math> medzi prúdom a napätím je v nastavenom rozmedzí</li></ul>
zemná nesmerová ochrana	$3I_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>3I_0</math></li></ul>
napät'ová ochrana	$3U_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>3U_0</math></li></ul>
admitančná nesmerová ochrana	$3I_0, 3U_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>Y_0</math></li></ul>
konduktančná smerová ochrana	$3I_0, 3U_0, \varphi_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>G_0</math></li><li>▪ uhol <math>\varphi_0</math> medzi prúdom a napätím je v nastavenom rozmedzí</li></ul>
konduktančná nesmerová ochrana	$3I_0, 3U_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>G_0</math></li></ul>
susceptančná smerová ochrana	$3I_0, 3U_0, \varphi_0$	<ul style="list-style-type: none"><li>▪ prekročenie nastavenej hodnoty <math>B_0</math></li><li>▪ uhol <math>\varphi_0</math> medzi prúdom a napätím je v nastavenom rozmedzí</li></ul>

Kombinácie jednotlivých ochranných funkcií (uvedené v Tab. 1) závisia od toho, či sú použité pre chránenie káblových, alebo vzdušných vedení a taktiež od spôsobu uzemnenia neutrálneho bodu siete [3], [4].

## II. PARAMETRIZÁCIA A OVERENIE ČINNOSTI ZEMNÝCH OCHRÁN

Z klasických ochrán boli pre testovanie použité:

- GSC 11 – generátorová alebo zemná, smerovaná a wattová ochrana
- GSCT 12X1 – generátorová alebo zemná, smerovaná, wattová a časovaná ochrana
- GSST 12X1 – generátorová alebo zemná, smerovaná, jalová a časovaná ochrana

Z digitálnych ochrán boli pre testovanie použité:

- RYo – multifunkčná zemná ochrana pre chránenie vedení
- MiCOM P139 – nadprúdová ochrana na ochranu a riadenie poľa
- SIPROTEC 7SJ632 – multifunkčná ochrana pre vývody a prenosové vedenia

Parametrizované boli zemné prúdové články. Aby bolo možné jednotlivé ochrany navzájom porovnať, ich parametrizácia je prispôbena možnostiam ochrany GSC a obmedzeniam ochrany SIPROTEC.

Parametrizácia v tomto prípade znamená nastavenie citlivosti zemného prúdového článku ochrán na detekciu prúdu  $I_0$  pre dva stupne citlivosti:

- 1 %  $I_N$
- 2 %  $I_N$

kde  $I_N = 5$  A pre všetky ochrany. Prúdový prevod PTP je na ochránach nastavený hodnotou 100/5 A.

Prúdy 1 %  $I_N$  a 2 %  $I_N$  sú menovité vypínacie prúdy ochrany, na ktoré ochrany nabiehajú. Pre 1 % je  $I_{0N} = 50$  mA a pre 2 % je  $I_{0N} = 100$  mA. Pre získavanie hodnôt týchto prúdov bol použitý binárny vstup testovacieho zariadenia CMC „B1“, ktorý je znázornený na schéme zapojenia ochrany RYo (Obr. 1).

Pri digitálnych ochránach dnes už odpadové prúdy nemajú veľký význam. V praxi sa ešte merajú z dôvodu dodatočnej informácie. Predstavujú 85 – 95 % rozbehovej hodnoty. Pre toto meranie boli odpadové prúdy konštantne znížené o 5 mA pre všetky ochrany a ich citlivosti. Pre získavanie hodnoty odpadových prúdov bol použitý binárny vstup zariadenia CMC „B2“, ktorý je znázornený na schéme zapojenia (Obr. 1).

Na testovanie citlivosti zemných prúdových článkov bol použitý softvér Test Universe s modulom *Ramping* [5]. Testom sa zistila rozbehová a odpadová hodnota prúdu.

Tabuľky nastavenia tzv. rampy (Tab. 2), na testovanie citlivosti prúdových zemných článkov všetkých ochrán, korešpondujú s tabuľkami v module *Ramping*. Nastavenie testu sa skladá z týchto hodnôt:

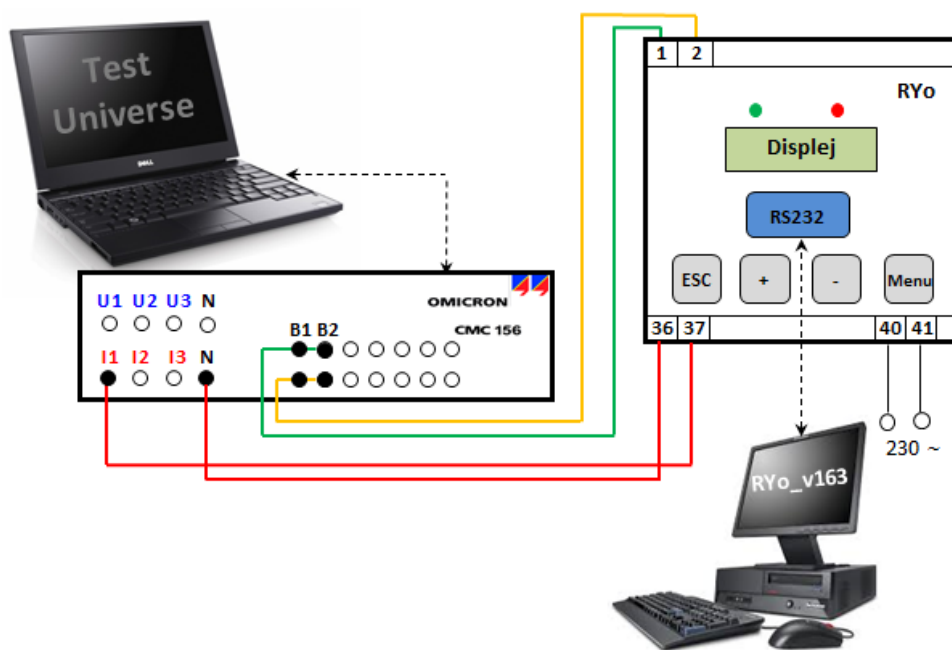
- „od“ – začiatok rampy
- „do“ – koniec rampy
- „delta“ – krok rampy
- „dt“ – čas zvýšenia / zníženia kroku rampy
- „I0“ – rozbehová / odpadová hodnota prúdu
- „Tol.+“ a „Tol.-“ – kladná a záporná odchýlka od rozbehovej / odpadovej hodnoty prúdu

Ďalším bodom parametrizácie bolo overenie časového pôsobenia prúdového článku. Na testovanie bol použitý modul *Quick CMC*. Parametrizácia spočívala v nastavení prúdu  $I_0$  na hodnotu  $1,5 I_N$  a overenie času pôsobenia ochrany pre časy  $t = 0$  s a  $t = 1$  s.

### A. Testovanie digitálnej ochrany RYo

Digitálna zemná ochrana RYo je určená na chránenie proti zemným spojeniam. Parametrizácia tejto ochrany bola realizovaná pomocou softvéru RYo\_v163 na PC. Pripojenie k ochrane bolo uskutočnené pomocou sériového portu RS232, ako je to znázornené na schéme zapojenia (Obr. 1). Parametrizovaná bola ochranná funkcia I0>, ktorú bolo potrebné pred použitím aktivovať. K aktívnej ochrane bolo priradené Relé 1. Následne bolo potrebné pomocou tohto softvéru parametre do ochrany zapísať.

Ochrana RYo na prevádzku potrebuje pomocné napájanie  $U_{\sim} = 230$  V.



Obr. 1 Schéma zapojenia pre testovanie ochrany RYo

Tab. 2 Nastavenie stavov testovacej rampy pre RYo

Nastavenie rampy 1 % $I_N$								
Signál $I_0$								
Rampa	od	do	delta	dt	d/dt	Kroky	Čas rampy	Stop podmienka
Stav 1	0 mA	60 mA	5 mA	1 s	5 mA/s	13	13 s	Nábeh 0→1
Stav 2	60 mA	0 mA	- 5 mA	1 s	- 5 mA/s	13	13 s	Nábeh 1→0
Nastavenie rampy 2 % $I_N$								
Signál $I_0$								
Rampa	od	do	delta	dt	d/dt	Kroky	Čas rampy	Stop podmienka
Stav 1	0 mA	110 mA	5 mA	1 s	5 mA/s	23	23 s	Nábeh 0→1
Stav 2	110 mA	0 mA	- 5 mA	1 s	- 5 mA/s	23	23 s	Nábeh 1→0

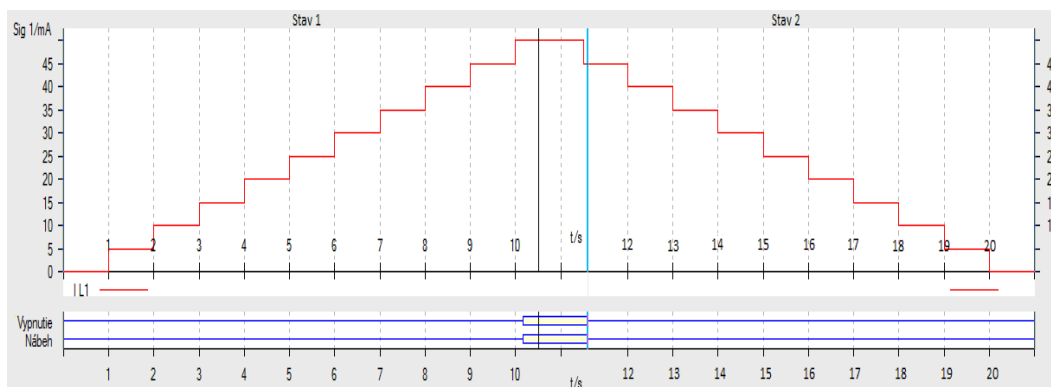
Tab. 3 Meranie rampy a výsledky pôsobenia ochrany RYo

Meranie rampy 1 % $I_N$							Výsledok testu			
Názov	Rampa	Podmienka	Signál	$I_0$	Tol.-	Tol.+	Akt.	Odch.	Vyhod.	Takt
Nábeh	Stav 1	Nábeh 0→1	$I_0$	50 mA	5 mA	5 mA	50 mA	0 mA	✓	172,3 ms
Odpad	Stav 2	Nábeh 1→0	$I_0$	45 mA	5 mA	5 mA	45 mA	0 mA	✓	92,3 ms
Meranie rampy 2 % $I_N$							Výsledok testu			
Názov	Rampa	Podmienka	Signál	$I_0$	Tol.-	Tol.+	Akt.	Odch.	Vyhod.	Takt
Nábeh	Stav 1	Nábeh 0→1	$I_0$	100 mA	5 mA	5 mA	100 mA	0 mA	✓	151,1 ms
Odpad	Stav 2	Nábeh 1→0	$I_0$	95 mA	5 mA	5 mA	95 mA	0 mA	✓	84,9 ms

**Výsledok testu: Test úspešný**

Výsledok testu bol úspešný, ochrana zodpovedá presnostiam uvedeným v katalógu výrobcu (Tab. 3).

Na nasledujúcom obrázku je možné vidieť postupný nárast rampy „Stav1“ až do okamihu kým na ochrane nastane popud na prúd  $I_0$ . Zariadenie CMC popud ochrany zaznamená a zastaví ďalší nárast rampy. Nárast rampy v tomto okamihu vystrieda pokles rampy „Stav2“, kým na ochrane neodpadne relé. Zariadenie CMC zaznamená odpad relé a ďalší pokles prúdu neuskutočňuje. Meranie ochrany sa v tomto okamihu končí a zariadenie CMC pomocou softvéru Test Universe vyhodnotí pôsobenie ochrany.



Obr. 2 Meraná rampa a pôsobenie ochrany RYo

Digitálna zemná ochrana RYo je časovaná, preto bolo pôsobenie ochrany nastavené okamžite  $t = 0$  s (resp. 0,1 s) a s časovým oneskorením  $t = 1$  s (Tab. 4). Ochrana testu vyhovela podľa tolerancie časového článku  $\Delta t = +100$  ms, ktorú udáva výrobca v katalógu ochrany.

Tab. 4 Nastavenie a výsledok časového pôsobenia ochrany RYo

Nastavenie časového pôsobenia pre 1 % $I_N$							Výsledok testu		
Názov	$I_0$	Štart	Stop	$T_{nom}$	Todch.-	Todch.+	Takt.	Todch.	Vyhod.
t pôsobenia	75 mA	$I_0$	Bin. Vstup 1 0→1	0,10 s	50 ms	100 ms	0,154 s	54 ms	✓
t pôsobenia	75 mA	$I_0$	Bin. Vstup 1 0→1	1,00 s	50 ms	100 ms	1,056 s	56 ms	✓
Nastavenie časového pôsobenia pre 2 % $I_N$							Výsledok testu		
Názov	$I_0$	Štart	Stop	$T_{nom}$	Todch.-	Todch.+	Takt.	Todch.	Vyhod.
t pôsobenia	150 mA	$I_0$	Bin. Vstup 1 0→1	0,10 s	50 ms	100 ms	0,180 s	80 ms	✓
t pôsobenia	150 mA	$I_0$	Bin. Vstup 1 0→1	1,00 s	50 ms	100 ms	1,083 s	83 ms	✓

Výsledok testu: **Test úspešný**

### III. ZÁVER

Príspevok ponúka oboznámenie sa s ochranami proti zemným spojeniam a demonštruje overenie ich činnosti. Podstatou je pochopenie princípov činnosti ochrán, ozrejmienie si problematiky zemných porúch. Získané informácie a skúsenosti sú nakoniec využité pre samotnú parametrizáciu a overenie činnosti týchto ochrán.

Ochrany, ktoré boli použité v tejto práci môžeme rozdeliť na dve skupiny:

- klasické – patria tu ochrany GSC 11, GSCT 12X1 a GSST 12X1
- digitálne – patria tu ochrany RYo, MiCOM P139 a SIPROTEC 7SJ632

Parametrizácia klasických ochrán proti zemným spojeniam spočíva v mechanickej konfigurácii na predných paneloch týchto ochrán. Digitálne ochrany boli parametrizované pomocou podporných softvérov. Ich konfigurácia je podstatne náročnejšia, ako pri elektromechanických ochránach.

Týmito parametrizáciami boli nastavené menovité vypínacie prúdy ochrán v dvoch stupňoch citlivosti 1 %  $I_N$  a 2 %  $I_N$ . Stupne citlivosti pre testovanie boli určené na základe možností ochrany SIPROTEC 7SJ632, ktorá nižší stupeň citlivosti nepodporuje.

Pre možnosť porovnania ochrán z hľadiska časového pôsobenia, bol na ochránach nastavený okamžitý čas pôsobenia  $t = 0$  s. Dôvodom bola ochrana GSC 11, ktorá časovaná nie je a jej pôsobenie je okamžité.

Sumár výsledkov testovania poukazuje na fakt, že elektromechanické ochrany vykazujú väčšie chyby citlivosti na nulovú zložku prúdu  $I_0$  ako digitálne ochrany. Časové pôsobenie elektromechanických ochrán bolo taktiež menej presné. Tieto chyby sú pravdepodobne spôsobené dlhodobou nečinnosťou ochrán a taktiež ich konštrukčným vyhotovením.

Digitálne ochrany vykazovali malé chyby zemných prúdových článkov. Časové pôsobenia ochrán sa pohybovali v predpísaných toleranciách, ktoré udávajú výrobcovia. Z toho vyplýva, že na detekciu zemných porúch je najvhodnejšie použiť digitálne ochrany. Túto skutočnosť možno potvrdiť na základe ich presnosti a rýchlosti pôsobenia.

Ochrany MiCOM P139 a SIPROTEC 7SJ632 sú digitálne terminály, ktoré obsahujú množstvo ochranných funkcií. Medzi ich funkcie patria aj také, ktoré sú určené na detekciu zemných spojení. Tieto terminály sú síce presné, ale ich všeobecné použitie zvyšuje ich cenu. Počet metód na detekciu zemných spojení je pritom obmedzený.

Ochrana RYo je zemná digitálna ochrana, ktorá je špecializovaná na zemné spojenia. Jej presnosť je porovnateľná s terminálmi MiCOM a SIPROTEC. Oproti týmto terminálom disponuje väčším množstvom metód na detekciu zemných spojení. Ochrana zahŕňa klasické princípy detekcie a je obohatená a princípy admitančné, konduktančné a susceptančné. Možnosť ich vzájomnej kombinácie umožňuje užívateľovi navoliť si nastavenie ochrany tak, ako si to vyžaduje daná konfigurácia siete. Veľkou výhodou tejto ochrany je aj možnosť zmeny metódy na detekciu zemných spojení počas prevádzky. Klasické ochrany takúto možnosť neposkytujú. Konštrukčné vyhotovenie týchto ochrán neumožňuje meniť ich princípy priamo v prevádzke. Ochrana RYo sa teda javí ako najvhodnejšie riešenie pre chránenie proti zemným spojeniam.

#### LITERATÚRA

- [1] *Zemní směrová relé GSC 11, GSCT 12X1 a GSST 12X1*. ZPA Trutnov.
- [2] F. Janiček, V. Chladný, A. Beláň, *Digitálne ochrany v elektrizačných sústavách*. TU Košice, 2003, 205 s., ISBN 80-89061-73-7.
- [3] *RYo – multifunkční zemní ochrana*. Protection & Consulting, Brno.
- [4] J. Pospíšil, M. Hvizdoš, "Nové přístupy k výuce problematiky zemních ochran v sítích VN," *6th International Scientific Symposium on Electrical Power Engineering – ELEKTROENERGETIKA 2011*, Stará Lesná, Sep. 2011, s. 234-237, ISBN 978-80-553-0724-4.
- [5] M. Hvizdoš, M. Kmec, *Skúšanie digitálnych ochrán pomocou testovacieho zariadenia CMC*. TU Košice, 2013, 98 s., ISBN 978-80-553-1507-2.

# Odporúčací Systém nad Databázou Filmov

*František BABIČ, Viktória BLANÁROVÁ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

frantisek.babic@tuke.sk, viktoriana.blanarova@student.tuke.sk

**Abstrakt** — Tento článok popisuje proces návrhu a implementácie odporúčacieho systému, ktorý umožní používateľovi postupne hodnotiť filmy prostredníctvom webovej aplikácie. Na základe dostatočného počtu odporúčaní mu budú následne odporúčané vhodné filmové tituly na základe podobnosti hodnotení iných používateľov. Ako zdrojové dáta boli použité existujúce databázy Freebase a MovieLens. Vhodnosť navrhutej metódy, využívajúcej princípy kolaboratívneho filtrovania, bola najprv testovaná voči klasickým predikčným metódam prostredníctvom podporného softvérového nástroja RapidMiner a následne bola použitá na implementáciu webovej aplikácie.

**Kľúčové slová** — odporúčací systém, RapidMiner, personalizácia

## I. ÚVOD

V súčasnej dobe je problematika odporúčacích systémov horúcou témou. Konzumenti prichádzajú dennodenne do styku s obrovským množstvom tovarov a služieb, kedy ťažko odhadujú ich relevantnosť. Na druhej strane sa firmy usilujú získať si a rozširovať základnú bázu zákazníkov. Jednou z príkladných oblastí využívania odporúčacích techník je aj filmový priemysel, ktorý predstavuje pomerne bohatý zdroj dát a rôzne možnosti ich využitia.

Existuje množstvo stránok, v rámci ktorých používatelia plávajú v mori informácií z dôvodu relatívne slabého navádzania vyhľadávacích nástrojov. A práve tento fakt motivoval vznik a vývoj nástrojov, pomocou ktorých by bolo možné ponúknuť používateľovi len relevantné informácie a tým si zabezpečiť jeho návrat. Jednou z možností v tomto prípade je personalizácia, čiže webová stránka nastavená tak, aby sa adaptovala individuálnym potrebám používateľa [1].

Základným cieľom personalizácie je teda odporúčanie takých produktov a služieb, ktoré sú pre používateľa alebo zákazníka relevantné aj s ohľadom na postupný rapidný nárast portfólií [2]. V súčasnosti však existuje široké množstvo rôznych techník personalizácie, ako napr. cookies, personalizácia založená na profile, osobné nástroje, oportunistické linky alebo odporúčacie systémy. Odporúčacie systémy poskytujú personalizované odporúčania založené na preferenciách používateľa. Pre tento typ personalizácie existuje viacero metód, medzi najpoužívanejšie patria kolaboratívne filtrovanie a filtrovanie založené na obsahu.

Burke vo svojej práci [3] uvádza kolaboratívne filtrovanie ako pravdepodobne najrozšírenejšiu, najpoužívanejšiu a najprepracovanejšiu z dostupných technológií pre personalizáciu. Systém získava hodnotenia alebo odporúčania objektov, rozpoznáva spoločné črty medzi používateľmi a na základe týchto podobností ich zaraďuje do samostatných zhlukov. Nové odporúčania sa generujú podľa preferencií používateľov v rámci zhľuku. Typický profil používateľa v kolaboratívnom systéme pozostáva z nepretržite dopĺňaného vektora. Názory používateľov sa môžu v priebehu času postupne meniť, preto niektoré systémy po určitej dobe hodnotenia vymazávajú. Medzi prvé odporúčacie systémy využívajúce techniku kolaboratívneho filtrovania patria napr. GroupLens<sup>1</sup> alebo Tapestry<sup>2</sup>. Ďalší zástupca s názvom Firefly<sup>3</sup> bol vytvorený výskumníkmi z MIT a neskôr ho kúpil Microsoft za približne 40

---

<sup>1</sup> <http://www.grouplens.org/>

<sup>2</sup> <http://www.cs.cmu.edu/~dmaltz/main-report/node10.html>

<sup>3</sup> [http://en.citizendium.org/wiki/Recommendation\\_system](http://en.citizendium.org/wiki/Recommendation_system)



miliónov dolárov.

Najväčšia výhoda kolaboratívnych techník je, že sú kompletne nezávislé od akejkoľvek reprezentácie pre počítač zrozumiteľných metadát odporúčaných objektov a sú tiež vhodné pre komplexné objekty ako napríklad hudba a filmy, ktorých rôznorodosť a štýl zodpovedá za veľké množstvo odlišností v preferenciách používateľov.

## II. NÁVRH A IMPLEMENTÁCIA SYSTÉMU

Proces návrhu a implementácie odporúčacieho systému bol realizovaný v súlade s metodológiou CRISP-DM, ktorá predstavuje všeobecný rámec pre riešenia úloh analýzy dát. V našom prípade boli jednotlivé fázy tejto metodológie reprezentované pochopením a analýzou dostupných dátových množín; predspracovaním a prípravou vybraných vzoriek záznamov z týchto databáz; experimentálne overenie vhodnosti vybraných metód personalizácie; vytvorenie výslednej webovej aplikácie v poslednej fáze CRISP-DM (nasadenie).

### A. Dátové Množiny

Náš výber bol orientovaný na voľne dostupné dátové množiny obsahujúce informácie o filmoch alebo hodnotenia filmov zo strany používateľov. Výsledná analýza bola zameraná na ich vhodnosť pre stanovenú úlohu, veľkosť a formát, v ktorom poskytujú dáta. V tejto sekcii stručne predstavíme niekoľko z nich. Freebase predstavuje rozsiahlu kolaboratívnu službu obsahujúcu metadáta podobnú Wiki, ktorá poskytuje otvorenú sémanticky značkovanosť databázu znalostí z rôznych domén a jej obsah môžu užívatelia editovať. Internet Movie Database (IMDb) je rozsiahla on-line databáza informácií súvisiacich s filmovými titulmi, televíznymi programami, videohrami. Je to jedna z najpopulárnejších stránok s touto tematikou, automaticky sa pokladá za najväčšiu svetovú databázu filmov. Linked Movie Database vznikla ako výsledok projektu z roku 2008, ktorý ponúka prvú otvorenú sémantickú webovú databázu venovanú informáciám zo sveta filmu. The MovieDB je voľne prístupná databáza filmov, ktorú upravujú jej používatelia. Tento projekt bol vytvorený v roku 2008 a postupom času sa zmenil na jednu z najväčších databáz filmov, ktoré sú upravované svojimi používateľmi.

V oblasti odporúčacích systémov takisto existuje viacero zaujímavých alternatív, pomocou ktorých je možné získať odporúčania v oblasti filmového priemyslu. MovieLens je odporúčací systém, ktorý ponúka používateľom zoznam filmových titulov, ktoré by ich mohli zaujímať. Odporúčací systém je založený na preferenciách konkrétneho používateľa a využíva kolaboratívne filtrovanie. Netflix je najznámejšia americká on-line videopožičovňa, ktorá poskytuje prístup k množstvu filmov a seriálov za mesačný poplatok. Yahoo! Movies predstavuje nekomerčnú rozsiahlu sieť kolekcie podrobných informácií zo sveta filmu, ktorá obsahuje aj personalizovaný odporúčací systém. Flixster je americká sociálna filmová sieť určená pre verejnosť, kde si môže používateľ po prihlásení pomocou účtu Flixster, Facebook alebo Google+ zisťovať informácie o konkrétnych filmových tituloch.

Po dôkladnej analýze sme vybrali ako vhodnú dátovú základňu kombináciu záznamov z databáz Freebase a MovieLens.

### B. Predspracovanie Dát

Predspracovanie vybraných dátových množín bolo realizované v softvérovom nástroji RapidMiner. Prácu s dátami komplikoval fakt, že obe dátové množiny boli pomerne rozsiahle a výrazne sa odlišovali. Viaceré operácie predspracovania bolo nutné optimalizovať z hľadiska výpočtovej náročnosti, v ďalšom texte budú popísané len tie najhlavnejšie.

Na úvod bolo nutné integrovať informácie z oboch dátových množín do jedného celku, kde v rámci tejto integrácie bolo nutné častokrát riešiť nekonzistentnosť zápisu tých istých údajov. Napr. v jednej databáze bol uložený *názov filmu* ako „The Matrix“ a v druhej ako „Matrix, The“. Takisto *názov filmu* v jednej z databáz obsahoval aj rok výroby, ktorý bolo nutné extrahovať a uložiť do separátneho atribútu, t.j. *dátum výroby*. Kľúčovým atribútom pre spájanie záznamov bol následne očistený *názov filmu*, kde bolo nutné ešte vyriešiť situáciu výskytu viacerých filmov s rovnakým názvom pomocou atribútu *dátum výroby*.

V ďalšom kroku bolo riešené atribúty, ktoré v spoločnom popise obsahovali viacero hodnôt, ktoré bolo nutné separovať a zaradiť do samostatných atribútov, napr. mená hercov, režisérov, atď. Ku každému menu bola zároveň pridaná hodnota, či sa v danom filme vyskytuje alebo nie.

Nakoniec boli jedinečné ID herca nahradené jeho skutočným menom z dátovej množiny Freebase. Ukážka predspracovaných záznamov je na Obr.1.

Row No.	id	name	genre	tag	imdb	staring	directed_by	year
1	6073	Victim	Crime Dram	less than 30	http://www.in	Derren Nesbitt Norman Bird A	Basil Dearden	1961
2	2728	Spartacus	Action Dram	crucifixion R	http://www.in	Nina Foch Charles McGraw W	Anthony Mann Star	1960
3	3088	Harvey	Comedy	classical Os	http://www.in	Victoria Horne Fess Parker Je	Henry Koster	1950

Obr. 1 Ukážka výslednej dátovej množiny po aplikovaní všetkých operácií predspracovania.

### C. Experimenty

Experimenty boli realizované na základe troch vybraných metód: kolaboratívne filtrovanie, filtrovanie na základe obsahu a vlastná metóda. Ako referenčnú vzorku bola použitá náhodne vybraná dátová množina, ktorá obsahovala 100 záznamov z databázy MovieLens. Táto vzorka obsahovala ID používateľov, ID filmu a hodnotenie, ktoré tomuto filmu prideliť. Proces experimentálneho overenia bol vedený v súlade s postupom, ktorý je popísaný v článku Sadovského a Chena [4].

V tomto prípade v skutočnosti nejde o tvorbu odporúčaní, ktoré by sa najviac páčili používateľovi, ale o predikciu hodnotení, ktoré by používateľ prideliť danému filmu. Ide teda o dve rôzne úlohy; kým v prvom prípade ide iba o odporúčanie filmov, ktoré by sa používateľovi čo najviac páčili a nejde o filmy, ktoré by sa mu nepáčili; v druhom prípade ide aj o predikciu hodnotenia filmov, ktoré by sa používateľovi možno až tak nepáčili. Kým prvý postup je založený na odporúčaní najviac zaujímavých filmov pre používateľa, druhý postup je zameraný na predikciu hodnotenia akéhokoľvek filmu hodnoteného používateľom. V našom prípade sme postupovali druhým možným spôsobom.

Vstupná dátová množina bola rozdelená na tréningovú a testovaciu vzorku, na určenie presnosti jednotlivých metód sme vybrali strednú kvadratickú chybu – Root Mean Squared Error (RMSE) a relatívnu kvadratickú chybu - Root Relative Squared Error (RRSE), ktoré sú bežne používanými metrikami pre podobné typy úloh.

V prípade filtrovania založenom na obsahu bola najprv vytvorená matica údajov, ktorej riadky reprezentujú filmy a stĺpce predstavujú jednotlivé vlastnosti týchto filmov (herci, režiséri...). Vlastnosti filmov boli získané vo fáze predspracovania; cieľovou premennou predstavoval atribút *rating*. Predikcia bola realizovaná pre film, ktorý používateľ ešte nevidel na základe jeho predchádzajúcich hodnotení iných filmov. Ukážka výsledku z nástroja RapidMiner je na Obr.2.

Row No.	rating	art1	art100	art102	art103	art109	art11	art113
1	3	0	0	0	0	0	0	0
2	2	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0

Obr. 2 Ukážka výslednej matice pre predikciu hodnotenia pomocou metódy filtrovania na základe obsahu.

V prípade kolaboratívneho filtrovania bol očakávaný model pre predikciu hodnotenia filmu vytváraný na základe hodnotení jednotlivých používateľov. Vstupom do tohto modelu boli hodnotenia filmov od všetkých používateľov, ktorí zároveň hodnotili aj film, pre ktorý chceme predikovať hodnotenie. Riadky v matici predstavujú záznamy jednotlivých používateľov a stĺpce predstavujú hodnotenia pre jednotlivé filmy. Takisto sú tieto informácie doplnené o premenné, či daný používateľ hodnotil daný film. Cieľovou premennou je atribút *rating*, čiže hodnotenie filmu, pre ktorý sa vytvára príslušný model, vid'. Obr.3

Row No.	user_id	rating	oldfilm_1	oldfilm_2	oldfilm_6	oldfilm_10	oldfilm_19	oldfilm_32
1	1880	2	5	3	4	4	3	4
2	2080	3	0	0	0	0	0	0
3	2566	3	3	3	0	4	0	3

Obr. 3 Ukážka výslednej matice používateľov a ich hodnotení na základe metódy kolaboratívneho filtrovania.

Vlastná metóda na generovanie odporúčaní je založená na kombinácii princípov kolaboratívneho filtrovania a hľadania podobnosti medzi používateľmi navzájom na základe ich hodnotení filmov. Na začiatku bol vytvorený zoznam používateľov, ktorí hodnotili referenčný film, čiže bolo známe jeho hodnotenie, ktoré bude neskôr použité na overenie presnosti

vytvoreného modelu. Následne bol tento zoznam zoradený podľa počtu hodnotených filmov, ktoré boli hodnotené aj cieľovým používateľom. Z tohto zoznamu bolo vybratých 400 top používateľov, v prípade ktorých bola následne vypočítaná podobnosť voči cieľovému používateľovi podobným princípom ako v prípade kolaboratívneho filtrovania. Výsledná matica však bola oveľa hustejšia ako v predchádzajúcom experimente. V ďalšom kroku bol určená miera rozdielnosti medzi hodnoteniami cieľového používateľa a ostatných používateľov v prípade spoločne hodnotených filmov. Na základe tejto miery bolo vybraných 20 používateľov s najnižšou hodnotou rozdielnosti, to znamená s najvyššou mierou podobnosti. Po normalizácii na interval (0,5,1) boli tieto hodnoty použité ako váhy pre vážený priemer hodnotení filmu, hodnotenie ktorého chceme predikovať pre cieľového používateľa. Pre každého používateľa boli následne vyselektované filmy, ktoré získali hodnotenia viac ako 3 na stupnici od 1 (najhoršie hodnotenie) do 5 (najlepšie hodnotenie). Vypočítaný vážený priemer pre jednotlivé filmy určil mieru ich relevancie, pomocou ktorej bolo finálne odporúčané hodnotenie daného filmu pre cieľového používateľa (*newRating*), vid'. Obr.4.

Row No.	film_id	relevancy	count(newRating)
1	410	1.986	7
2	661	2.008	6
3	100	2.012	7

Obr. 4 Ukážka výsledného odporúčania pomocou vlastnej metódy.

#### D. Vyhodnotenie

Na vyhodnotenie uskutočnených experimentov bola použitá dátová vzorka 100 hodnotení. Tri použité metódy boli porovnané voči dvom referenčným, čiže jednoduchá voľba náhodnej predikcie hodnotenia filmu (Random) a priemer hodnotení všetkých filmov (Average). Výsledok porovnania je zobrazený v Tab.1.

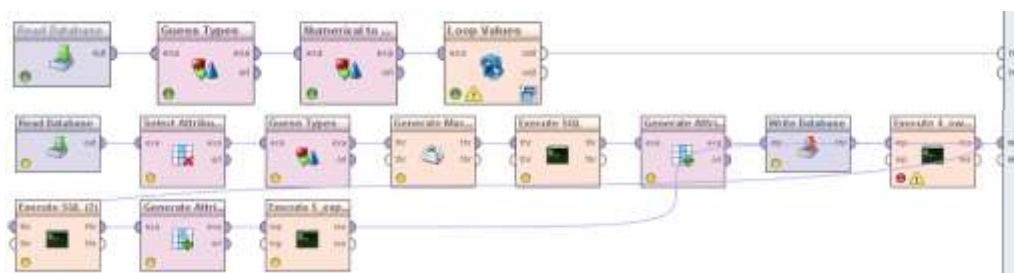
Tabuľka 1  
 Vyhodnotenie experimentov

	Kolaboratívne filtrovanie	Filtrovanie založené na obsahu	Vlastná metóda	Random	Average
RMSE	0.814	1.025	0.894	2.255	1.21
RRSE	1.062	1.337	0.745	1.88	1.009

#### E. Nasadenie

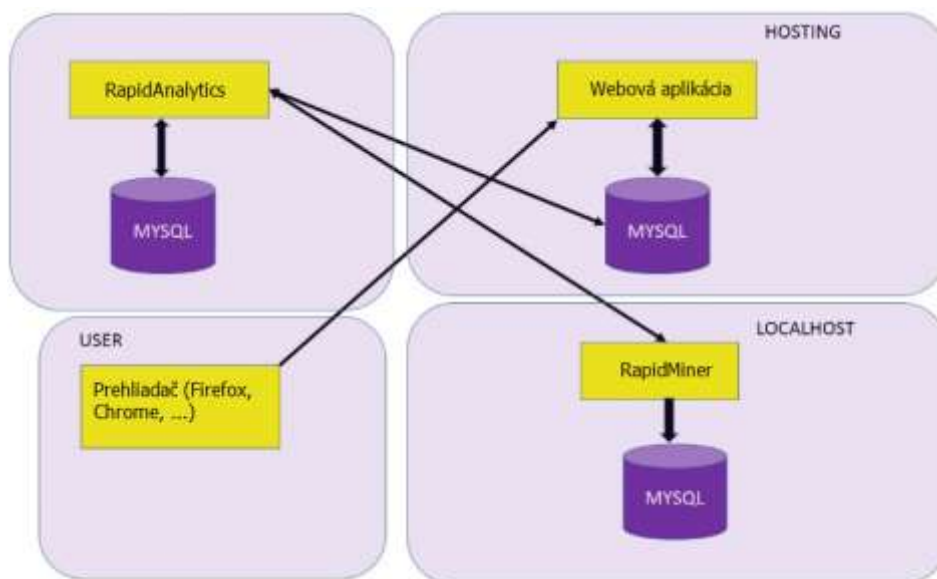
Vo fáze nasadenia bola použitá vlastná odporúčacia metóda na implementáciu odporúčacieho systému nad databázou filmov vo forme webovej aplikácie. Tento výber bol podložený najlepšou predikčnou schopnosťou a relatívne nízkymi nárokmi na výpočtovú kapacitu, ktoré sú napríklad u kolaboratívneho filtrovania pomerne vysoké. V rámci tejto aplikácie má používateľ možnosť, na základe filmov, ktoré hodnotil, získať odporúčané hodnotenia pre ďalšie filmy.

V prípade aplikačného servera RapidAnalytics sa proces zodpovedný za generovanie odporúčaní spúšťa opakovaním v určitých intervaloch každý deň, aby po overení, ktorí používatelia už majú zadaných viac ako 20 hodnotení, tak pre nich začne generovať odporúčanie pre ostatné filmy v databáze.



Obr. 4 Ukážka procesu, ktorý zabezpečuje generovanie odporúčaní.

Architektúra samotného systému je na Obr.5.



Obr. 5 Ukážka architektúry výsledného odporúčacieho systému nad databázou filmov.

### III. ZÁVER

Cieľom tejto práce bolo experimentálne overiť vhodnosť vybraných metód personalizácie pre vytvorenie odporúčacieho systému nad databázou filmov. Dosiahnuté výsledky experimentov potvrdili vyššiu presnosť predikcie u vybraných troch metód voči náhodnej stratégii alebo stratégii využívajúcej priemerné hodnotenie filmov. Z pohľadu výpočtovej náročnosti bola následne vlastná metóda použitá na implementáciu systému vo forme webovej aplikácie využívajúcej webové analytické služby ponúkané softvérovým nástrojom RapidMiner.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] D. Wu, I. Im, M. Tremaine, K. Instone, and M. Turoff, "A framework for classifying personalization scheme used on e-commerce Websites," in *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, 2003.
- [2] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, "Recommender systems," *Phys. Rep.*, vol. 519, no. 1, pp. 1–49, Oct. 2012.
- [3] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Model. User-Adapt. Interact.*, vol. 12, no. 4, pp. 331–370, Nov. 2002.
- [4] A. Sadovsky, X. Chen, "Evaluating the effectiveness of regularized logistic regression for the Netflix movie rating prediction task," p. 6, 2006.

# Metódy strojového učenia v analýze názorov

<sup>1</sup>Lubomír BAREJ, <sup>2</sup>Kristína MACHOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>lubomir.barej@student.tuke.sk, <sup>2</sup>kristina.machova@tuke.sk

**Abstrakt** — Tento príspevok sa venuje problému analýzy názorov a postojov, presnejšie klasifikácii jednotlivých diskusných príspevkov do konkrétnych tried na základe konverzačného obsahu. Rozbor týchto názorov ako aj získavanie informácií a znalostí z nich je pre človeka v dnešnej dobe nesmierne dôležité. Tieto nadobudnuté údaje môžu mať ľubovoľnú formu (recenzie, komentáre, rady, analýzy, apod.) a dajú sa zužitkovať skoro všade, napríklad pri kúpe produktu alebo služby, pri internetovom obchode, pri výbere filmu, v politickej alebo spravodajskej oblasti, apod. Na začiatku sa uvádza stručný opis problematiky, o ktorej riešenie sa práca pokúša. Ďalšia časť sa venuje návrhu vlastného riešenia daného problému a jeho implementácii. V závere sú opísané vykonané experimenty danej implementácie s využitím testovacej množiny nezávislých príspevkov ako aj spôsob výpočtu presnosti a návratnosti hodnotených výsledkov.

**Kľúčové slová** — strojové učenie, Naivný Bayes klasifikátor, analýza sentimentu, klasifikácia názorov

## I. ÚVOD

Významnou súčasťou ľudského chovania bolo vždy mať určitý názor na danú vec, ale ešte dôležitejšie bola pre nich túžba zistiť, čo si myslia o tom ostatní ľudia. S rastúcou dostupnosťou a popularitou informačných technológií sa prostredníctvom internetu, web stránok, online blogov apod. objavujú nové možnosti a výzvy využívania týchto zdrojov, obsahujúcich obrovské množstvo názorov a postojov na prehládavanie a pochopenie názorov druhých.

Manuálne hľadanie, zisťovanie a spracovanie týchto informácií je dosť zložité, časovo príliš náročné a nekomfortné. Preto sa na odstránenie týchto problémov začína používať umelá inteligencia, ktorá zažíva v poslednej dobe veľký rozmach, a preto sa stáva bežnou súčasťou ľudského života.

Cieľom v tomto príspevku je vybrať a následne použiť vhodnú metódu umelej inteligencie na riešenie problému analýzy názorov. Presnejšie ide o problém klasifikácie príspevkov do tried: pozitívne a negatívne príspevky na základe predchádzajúceho ohodnotenia jednotlivých slov v danom článku z hľadiska pozitívnosti. Zo všetkých metód strojového učenia bol zvolený *Naivný Bayesov klasifikátor*, ktorý sa osvedčil pri spracovávaní textu.

## II. METÓDA NAIVNÉHO BAYESOVHO KLASIFIKÁTORA

Vrámcí strojového učenia sa Naivný Bayesov klasifikátor zaraďuje do skupiny algoritmov na kontrolované učenie resp. učenie s učiteľom, čo znamená, že sa predpokladá, že výstup je vopred známy. Táto metóda je veľmi často využívaná najmä kvôli jeho flexibilitě, jednoduchosti a zrozumiteľnosti. Naivný Bayesov klasifikátor je často používaný v oblastiach ako sú:

- dolovanie dát - spôsob ako získať skryté informácie, vzťahy medzi nimi alebo určité vzory,
- klasifikácia textu - prostredníctvom analýzy textu z rôznych perspektív a jej súhrnom vie Naivný Bayesov klasifikátor získať užitočné informácie, ktoré sa ďalej môžu využiť,
- rozpoznávanie názorov - pomocou extrakcie a vyhodnotenia dokáže Naivný Bayesov klasifikátor určiť názor jednotlivých používateľov, ktorí tieto príspevky, komentáre alebo texty pridali na web,
- spam filter - táto metóda patrí medzi efektívne systémy na rozoznávanie nevyžiadanej pošty, nevhodných príspevkoch na webe, apod.,

- odporúčací systém - na základe skrytých informácií v texte, dokáže systém získať a predvídať či je daný zdroj pre používateľa vhodný.

Patrí medzi najstaršie používané algoritmy, ale aj napriek tomu je to stále jedna z najpoužívanejších metód v rôznych oblastiach, v poslednom čase hlavne v oblasti, ktorej je venovaný tento príspevok, t.j. klasifikácia textu a analýzu názorov. V podstate je to pravdepodobnostný klasifikátor založený na aplikácii Bayesovej teóremy:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Kde,

$P(A|B)$  = pravdepodobnosť splnenia javu A za podmienky, že je splnený jav B,

$P(B|A)$  = pravdepodobnosť splnenia javu B za podmienky, že je splnený jav A,

$P(A)$  = pravdepodobnosť splnenia javu A,

$P(B)$  = pravdepodobnosť splnenia javu B.

Princíp tejto metódy spočíva v tom, že sa nedá jednoznačne povedať, do ktorej triedy príklad, ktorý vyšetrujeme patrí, pretože v podstate môžeme povedať, že úloha patrí do všetkých tried s určitou pravdepodobnosťou. Cieľom algoritmu je nájsť len tú triedu, ktorá má najvyššiu pravdepodobnosť. Naivita Bayesovho klasifikátora spočíva v tom, že predpokladá nezávislosť jednotlivých atribútov voči sebe. To znamená, že efekt hodnôt každého atribútu na danú triedu nie je ovplyvnený hodnotami iných atribútov.

### III. PRÍPRAVA TRÉNOVACEJ MNOŽINY

V prvom rade bolo potrebné zozbierať a predspracovať množinou dát na testovanie prístupu ku klasifikácii príspevkov z webových diskusií jednotlivých názorových tried. Po hlbšej analýze sa zistilo, že dané príspevky neboli nezávislé, a preto by pre riešenie daného problému boli nevhodné. Hlavnou príčinou toho bolo to, že používatelia namiesto zaslania príspevku, v ktorom by uviedli svoj kladný alebo záporný názor k danej téme, skôr zverejňovali príspevky, v ktorých uvádzali vlastné postoje, ktoré nesúviseli celkom s danou témou. Preto sme sa rozhodli venovať skôr web stránkam, ktoré poskytujú diskusie, komentáre a hodnotenia na konkrétne produkty a služby, ako napríklad spotrebný tovar, filmografia, ubytovacie zariadenia, apod. Najviac dát sa získalo z webových stránok [www.csfd.cz](http://www.csfd.cz), ktorá sa zaoberá celosvetovou filmografiou (zoznamy filmov, hodnotenia, komentáre, apod.) a internetovým obchodom [www.alza.sk](http://www.alza.sk), ktorý poskytuje predaj rôzneho sortimentu výrobkov (ku každému výrobku sa dajú pridať pozitíva, negatíva, zhrnutie prostredníctvom krátkeho popisu daného tovaru a spokojnosť v podaní percentuálneho ohodnotenia 0 - 100 [%]).

Následne sa získané dáta modifikovali a upravili na taký tvar, aby bolo možné s nimi vo vytvorenom programe pracovať. Zbieranie vhodných údajov bolo pri tom ovplyvnené vopred určenými kritériami, ktoré mali zabezpečiť efektívne a optimálne fungovanie systému.

Boli vytvorené tri textové súbory:

1. *súbor* - obsahuje jednotlivé položky, ktoré tvoria bázu dát (pozitívne a negatívne slová). Pre lepšiu a rýchlejšiu prácu môjho programu pri detekovaní daných slov boli upravené na základný tvar slova.
2. *súbor* - obsahuje početnosť konkrétnych slov z predchádzajúceho súboru v príspevkoch, ktoré boli použité pri tvorbe tejto trénovacej množiny.
3. *súbor* - obsahuje polaritu jednotlivých položiek ("+" pozitívna polarita, "-" negatívna polarita).

Veľkosť každého z daných troch súborov je 2017 položiek. Tieto súbory sa pri používaní programu spájajú do jedného celku, pomocou ktorého sa vyhodnocuje príslušnosť príspevku do konkrétnej triedy, ktorá sa na výstupe vypíše.

### IV. KLASIFIKÁCIA NÁZOROV

Prezentovaný prístup ku klasifikácii názorov pozostáva z nasledujúcich krokov:

#### A. vkladanie textu

Po spustení aplikácie sa od používateľa očakáva vloženie príspevku, ktorý sa má analyzovať. Môžu nastať dva prípady:

1. Vloženie textu prostredníctvom .csv súboru - pri výbere tohto spôsobu bude skrátený čas spracovania, keďže sa všetky príspevky v danom súbore spracovávajú paralelne. Pri tomto spôsobe je vysoko pravdepodobné, že na vstupe program dostane väčšie množstvo príspevkov, preto bol zvolený formát .csv súbor, lebo umožňuje články medzi sebou rozlišovať pomocou zvoleného znamienka (v tomto prípade ";"). Program

používateľovi nedovolí vybrať súbor, ktorý je iného typu, pretože v okne, ktoré sa mu otvorí za účelom výberu nového dokumentu, uvidí len súbory v danom špecifickom formáte vďaka filtru, ktorý je v programe nastavený.

2. Vloženie textu manuálne - pri tomto spôsobe používateľ vkladá len jeden príspevok. Avšak aplikácia umožňuje nastavenie pri ktorom je možné vkladať viac článkov za sebou, a tak je možnosť ohodnotiť objektívne celú diskusiu o danom objekte.

#### B. *predspracovanie dát*

V tomto bode bolo potrebné riešiť problémy súvisiace so spracovaním prirodzeného jazyka. Bolo potrebné ošetriť tri prípady:

1. diakritické znamienka - pri vkladaní príspevkov do programu bolo zavedené obmedzenie, ktoré upozorňuje používateľa, aby vkladal texty bez veľkých písmen a diakritiky,
2. interpunkcia - interpunkčné znamienka boli odstránené pri delení textu na slová,
3. odstránenie určitých slov - bolo zabezpečené, aby po vložení textu boli odstránené všetky neplnovýznamové a pomocné slová.

#### C. *aplikácia umelej inteligencie*

Po modifikácii jednotlivých slov z vyšetřovaného príspevku sa pristúpi k použitiu Naivného Bayesovho klasifikátora. Následne sa každé slovo z článku porovná s každým slovom v báze dát.

Pri porovnávaní slov sa používa *regulárny výraz* (skr. *regex*), ktorý sa automaticky aplikuje na každé jedno slovo z bázy dát, ktoré sa porovnáva. *Regex* je v podstate sekvencia znakov, ktoré umožňujú prácu s textom na základe určitých syntaktických pravidiel a je súčasťou mnohých programovacích jazykov, ako aj textových editorov v Unixe. V tomto prípade sa *regex* využíva na to, aby doplnil chýbajúce písmená a znaky ku každému slovu z bázy dát a našiel zhodu so slovom zo skúmaného príspevku. Ak sa slová zhodujú, tak sa následne slovu priradí váha, ktorá je vypočítaná nasledovne:

$$\text{váha} = \frac{\text{počet výskytov konkrétneho slova v danej triede}}{\text{počet príspevkov v danej triede}}$$

Po výpočte váh jednotlivých slov sa určí pravdepodobnostná hodnota príspevku, ktorá zostáva zo súčtu váh jednotlivých slov, ktoré sa zhodli so slovom z bázy dát. Následne sa výsledná hodnota porovná, či je pozitívna alebo negatívna a zaradí sa do danej triedy polarity.

### V. TESTOVANIE IMPLEMENTÁCIE

Navrhnutý prístup ku klasifikácii názorov bol implementovaný a testovaný. Prostredníctvom týchto testov sa demonštruje funkčnosť a efektívnosť programu. Boli vykonané tri nezávislé testy, pričom sa postupne zlepšovala báza dát.

Konečné hodnotenie efektívnosti programu sa posudzovalo na základe dvoch parametrov, a to presnosti ( $\pi$ ) a návratnosti ( $\rho$ ) podľa nasledovných formlí:

$$\pi_j = \frac{TP_j}{TP_j + FP_j} \qquad \rho_j = \frac{TP_j}{TP_j + FN_j}$$

Kde,

$TP$  = počet príspevkov, ktoré boli označené ako pozitívne a v skutočnosti sú pozitívne,

$FP$  = počet príspevkov, ktoré boli označené ako pozitívne a v skutočnosti sú negatívne,

$FN$  = počet príspevkov, ktoré boli označené ako negatívne, ale v skutočnosti sú pozitívne.

V prvom experimente sa pracuje s testovacou množinou, ktorá obsahuje 100 príspevkov, z toho je približne 55 kladných a 45 záporných príspevkov. Program tu pracuje s pôvodnou bázou dát, ktorá bola vytvorená na začiatku. Po prvom testovaní sa zistilo, že vypočítané hodnoty pre negatívnu triedu sú nedostačujúce, preto sme sa v nasledujúcich experimentoch zamerali na postupnú modifikáciu bázy dát doplnením nových slov a početností. V rámci tretieho experimentu bola tiež modifikovaná báza dát, úpravou niektorých početností a zahrnutím nových príspevkov. Výsledné hodnoty presnosti a návratnosti pre jednotlivé testy sú uvedené v nasledujúcej tabuľke Tab.1.

Tabuľka 1  
Hodnoty presnosti a návratnosti troch rôznych experimentov.

	Pozitívna trieda		Negatívna trieda	
	Π [%]	P [%]	Π [%]	P [%]
test č. 1	62,6	83,9	64,2	36,4
test č. 2	65,7	85,7	70,4	43,2
test č. 3	73,0	75,6	67,1	65,8

## VI. ZÁVER

Cieľom tohto príspevku bolo podať stručný prehľad o problematike strojového učenia v rámci riešenia úlohy analýzy názorov v rámci konverzačného obsahu. Bol navrhnutý a implementovaný prístup k riešeniu problému klasifikácie príspevkov do dvoch tried polarity názorov.

Pri testovaní sa vykonali celkovo tri experimenty a z toho najlepšie výsledky dosahovala presnosť pri pozitívnej triede vyše 73% a pri negatívnej triede vyše 67%. Ukázalo sa, že hodnoty presnosti a návratnosti boli silno ovplyvnené aj malými modifikáciami v báze dát. Aby sa zlepšili hodnoty presnosti a návratnosti bude potrebné v budúcnosti zmeniť spôsob váhovania slov. Ďalší návrh na zlepšenie, ktorý by sa mohol realizovať je spracovanie viacerých slov v postupnosti za sebou použitím tzv. n-gramov. Bežným problémom pri klasifikácii slov je, že napríklad pozitívne slová sa používajú na vyjadrenie príspevkov s negatívnym názorom a naopak.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] E. Barej, „*Metódy strojového učenia v analýze názoro.*“ Bakalárska práca, Technická univerzita, Košice, 2014, 59s.
- [2] T. Dietterich, „Machine Learning“ [Online]. Dostupné na internete: <http://web.engr.oregonstate.edu/~tgd/publications/nature-ecs-machine-learning.pdf>
- [3] K. Chai, H. Chieu, „*Bayesian Online Classifiers for Text Classification and Filtering*“ [Online], 2002. Dostupné na internete: <http://software.ucv.ro/~cmihaescu/ro/teaching/AIR/docs/Lab4-NaiveBayes.pdf>
- [4] B. LIU, „*Sentiment Analysis and Opinion Mining*“ Morgan & Claypool Publishers, 2012



# Advanced organization of Bibliography and Citations

<sup>1</sup>*Liberios Vokorokos*, <sup>1</sup>*Branislav Madoš*, <sup>1</sup>*Marek Čajkovský*, <sup>1</sup>*Ján Hurtuk*, <sup>1</sup>*Juraj Rákoš*

<sup>1</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>liberios.vokorokos@tuke.sk, <sup>1</sup>branislav.mados@tuke.sk, <sup>1</sup>marek.cajkovsky@tuke.sk,  
<sup>1</sup>jan.hurtuk@tuke.sk, <sup>1</sup>juraj.rakos@student.tuke.sk

**Abstract** — Organization of the content of references is problematic when it has to be done by an algorithm. Computers are not designed to understand semantics and when they have to process information that relies on implied information it can lead to problems. This paper analyzes existing solutions to this problem and introduces a new solution for the organization of references.

**Keywords** — Organization, parsing, bibliography, reference, citation

## I. INTRODUCTION

Analyzing text and manly its meaning is a task that is not well suited for computers. While computers are able to do large amounts of work and process very large amounts of data in a very short amount of time due to their raw processing power they stumble on the analysis of text due to not understanding semantics behind it. Parsing a sentence while trivial for a human can be a very problematic task for a computer. Humans are way more better at understanding the semantics of the text. Natural language processing is a field that has been trying to make these kind of applications possible. In recent times more complex applications and methods were created as the result of these efforts. Some of these efforts have already come to fruition in the form of commercially available application that solve similar problems.[1] References are even harder compared to normal sentences. Higher amount of information in a reference is only implied and not explicitly stated. Humans can easily understand this implied information which is lost on an algorithm. This makes organizing references into a more machine readable format very hard. To do so it is necessary to create an algorithm that would try to solve these problems and which can then organize and connect the implied information like Publisher with its stated counterpart. Organizing references by hand would take both time and resources that can then be spent elsewhere. Creating a software solution which could do the same thing with the added advantage of how quickly an algorithm can wade through large swaths of data is thus very desirable. Over the years resources were poured into fields that could be used for problems that require a certain level of abstraction. This research was done in fields of natural language processing, learning algorithms and knowledge based systems. This research helped advance these fields to a point where many of these tasks have viable commercial solutions.

## II. ANALYSIS

The organization of references is complex topic due to several factors. The first is the fact information in a reference is very condensed thus making it hard to differentiate information. Algorithms run into problems once implied metadata needs to be taken into account to correctly process information. The creation and usage of references was an organic process which did not happen overnight. Due to this many organizations created guidelines on how to properly reference the work of others. These organizations often had differing goals in mind when they created their own guidelines causing a situation where there are many competing reference styles which can be used. These styles differ from each other based on the needs of the organization that created them. These differences can vary between the styles. This is another

factor that needs to be taken into account when working with references. The differences between these styles make a software solution for organization of references more complex to account for these differences. This is even more complicated by the fact that a reference itself does not contain the information which style was used to write it. This means that a program cannot simply look for the information on which style was used to know how to process the reference further. Thus the software has to account for the differences between references even without knowing which style was used to create reference. Due to the condensed nature of references it is not uncommon for human error to happen when writing a reference by hand. Errors like missing or miss placed commas and full stops also need to be addressed. Under normal circumstances these symbols might prove helpful when identifying parts of reference from each other. They provide a useful boundary between information that would help with parsing and organizing a reference. However this boundary should not be taken for granted. An error like this could cause an algorithm to miss attribute information. That would cause the program to incorrectly organize data. A robust software solution should be immune to this problem.

#### *A. Existing methods*

People have tried to solve this problem before. Several approaches were already used with varying degrees of success. The existing methods are probably best summed into three distinct categories which are:

- Regular expression based on heuristics and heuristic algorithms
- Learning algorithms
- Knowledge based systems

#### *B. Regular expression based heuristics and heuristic algorithms*

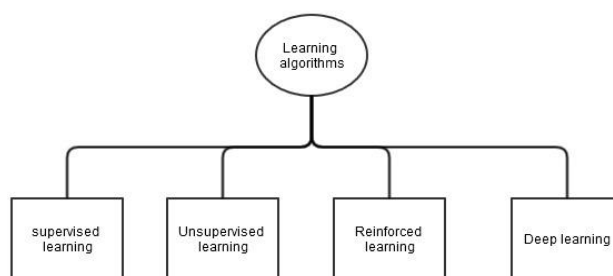
This method as the name implies relies heavily on regular expressions to parse and organize the reference. Using regular expressions is like using a static solution for a dynamic problem. When using regular expressions the program will always give the same result no matter how it is run. To achieve some difference in accuracy of the program it is necessary to create more complex regular expressions. Compared to other solutions however regular expressions are less complex because they do not change due to being ran. This can be both an advantage and a disadvantage. Other methods can learn from past experience and become more accurate. This also introduces the risk that the program might change in a way that lowers its accuracy in which case it would have been better if it had stayed the same. [3] An example of regular expression based heuristics is the paracite toolkit. [2] which is a collection of PERL modules which can be used to extract metadata from references. A disadvantage to regular expression based heuristics is the fact that a person must make the decision as to what the most likely scenario the program might encounter be. This means that the programmer must make decisions what functionality is important enough have a higher priority to be made. If the programmer chooses poorly it might lead to a program that has a very high accuracy but only in very specific circumstances that are very unlikely to happen. The programmer has to make these decisions as to what functionality to include including of many special rules for cases would quickly bloat the code and make it unmanageable. Also such conduct would quickly raise its cyclomatic complexity. They can be used in many different fields, like pattern matching etc. [11]

#### *C. Learning algorithms*

This method is based on a self-changing concept. They begin with a set of rules and some amount of training data. During its operation they make adjustments based on the data they encounter making them theoretically more accurate the more data they have at their disposal. Compared to regular expression heuristics where any changes must be made directly into their code this type of algorithms just need better training data.[4] An example of learning algorithms would be the hidden markov model.[5][10] Learning algorithms encompass two terms that are often confused. These are machine learning and data mining. Machine learning tries to predict new data based on known properties. Data mining focuses on finding new properties from unknown data. Both of these are necessary in a functional learning algorithm.[6] The accuracy in learning algorithms is highly reliant on the training data supplied to the program. Bad training data could lead to very low results. The training data however has to be compiled by a person. Future data is hard to predict and thus compiling a good set of training data is very hard.[4]

Learning algorithms are not perfect. Their training data are always finite and unknown future data are hard to predict making a "perfect" learning algorithm virtually impossible. Also for some types of problems there might be not enough training data available to teach the algorithm what it is supposed to be doing. There is the option of just making more training data but is a scenario where effort and gain has to be considered. Is it worth the effort creating large amounts of training data for a scenario that is very unlikely to occur? The answer would depend on a few factors like how big of an impact would this decision have on the accuracy of the algorithm and how many man-hours it would take to create this training data. [8] Learning algorithms can be divided into several categories by type. They are divided by the method of learning new information. These are[9]:

- Supervised learning
- Unsupervised learning
- Reinforced learning
- Deep learning



**Figure 1 Learning algorithms categories**

#### *D. Knowledge based systems*

Knowledge based systems which are also sometimes called expert systems tend to provide better results than alternatives but they tend to have a narrower scope. They use stored knowledge to solve problems. Knowledge based systems use a database of sorts. Historically, the database had very different requirements compared to normal databases. Knowledge bases only had one user so multi user access was not important and had not been taken as needed. Neither were integrity checks of the database. Because the knowledge based system had only one user at a time there was no need for integrity checks to prevent bad commits to the database. Security was also considered as a non issue because to use of the knowledge based system the user had to have physical access. Thus there were different requirements placed on knowledge based systems that made several features of regular databases unnecessary. However, over time the usage of knowledge based systems has been changed. They started to be used by more users and thus knowledge based systems had to change and be more like regular databases that are used in a corporate environment. Knowledge has to be represented in a way for the program to understand. Knowledge is represented with rule frames and scripts.[7] Knowledge based systems require three components to work properly and these are:

- Knowledge base
- Acquisition mechanism
- Inference mechanism

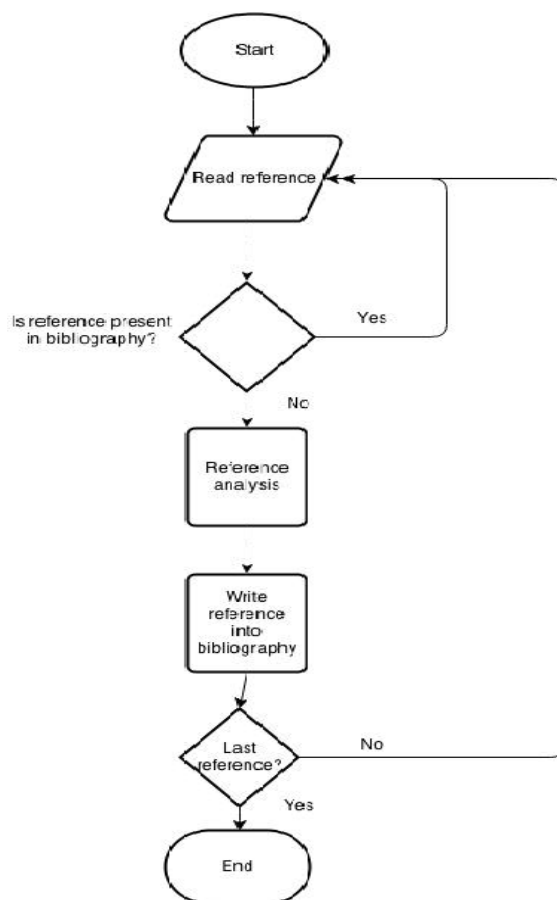
### III. SOLUTION MECHANISM

After analyzing the currently existing methods it is necessary to create a methodology to create a solution that would fulfill the necessary requirements. The first decision made was the choice of programming language. After much deliberation python was chosen as the best candidate for the task. Python has some advantages that make it ideal for this kind of work. Python is easy to

understand and learn, there is a high amount of tutorials and study materials from which python can be learned. Python does not need to compiled leading to a shorter iteration programming process but this feature also can be considered as a disadvantage because of the fact that if a part of program has not run, there will not be any errors referring it. The basic schema of the conceptual program proposal is shown on the next figure (figure 1). It shows the interaction of the three main components of the program which are :

- Bibliography comparison
- Reference analysis
- Addition of the reference into the bibliography

As seen on the figure (see figure 1 ) the program first checks whether the reference is already in the bibliography. If it is indeed present the program will continue on to the next reference until a new reference is found. When a new reference is found the program will start to analyze this reference based on preset rules. These rules include several dictionary files which are checked whether the currently analyzed word is on them. If, for example the currently analyzed word is on the name list the program will be more likely to assume that the word is part of the author.



**Figure 2 The program conceptual schema**

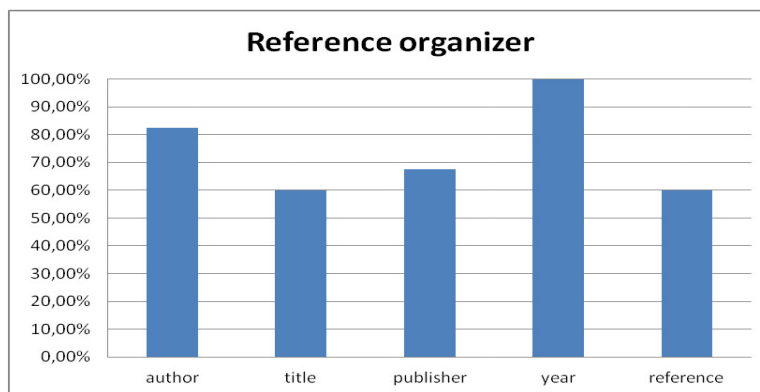
The operation of the program could be described as summary in these steps:

1. Split reference to words
2. Check reference against bibliography
3. Assign values to words
4. Assign words to categories
5. Write reference into bibliography

The program also supports cases where there are two authors listed under the same work. This is done by checking if the detected author contains a separator like and or & which serve the purpose of separating multiple authors in a reference. If a separator is present the program will split the author into two parts. Thus the program writes the authors individually and the separator will be left out from the entry.

#### IV. TEST OF THE RESULT

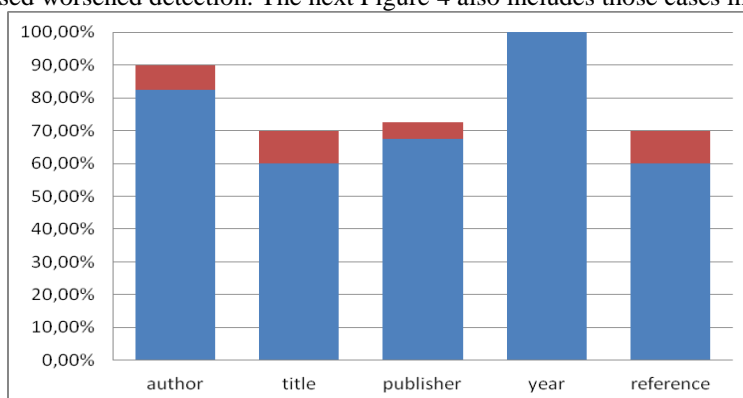
To verify the functionality of the solution it is necessary to conduct some tests to evaluate the program. For the testing a sample of references was chosen to be tested. The program then tried to correctly tries to organize the information provided in the reference into categories. The output is then evaluated. The output is then checked whether author, title, publisher and year are assigned correctly. Each of these is rated independently and at the end the overall performance is rated. Overall performance means only those cases where all the fields of the reference were correctly read without omissions or inaccuracies. The results of the testing can be seen in the figure (see figure 2 ) below.



**Figure 3 The graph scheme representing testing results**

The testing showed high accuracy of year and author detection. This is due to the fact that detecting a year is relatively easy and that the extensive dictionary file of names used for detection of names helps considerably. The detection of the title and publisher is harder comparing to other data. The relatively worse detection was caused by worse detection of the boundary between the title and publisher. If the algorithm thought the boundary is one word of it caused one field to be a word short and the other to have one word extra. This however meant that the result had to be marked wrong for both. This is why the detection rate of both is identical. Also the overall detection is identical to this value only proving that this boundary is the bottleneck for detection and improving it further would lead to yielding results.

There were cases where a reference was not correctly detected because one word got moved from one category to the other. These were for example surnames at the end of the author field. The last surname of the author field was not correctly detected and got moved into the title and thus this caused worsened detection. The next Figure 4 also includes those cases into the results.



**Figure 4 Errors made in the detection proces**

As can be seen from the figure the number of cases where the error in detection was only by one word is not insignificant. The blue part of the graph shows the same values as the previous graph. The red part shows the difference if the cases where one moved word caused bad detection. As it shows that the number of these cases is high enough to warrant attention.

## V. CONCLUSION

Organizing references is a very hard task if done manually. Automation of this task could save a lot of manpower and resources in the future. Several methods of trying to organize a reference into parts and add appropriate metadata have already emerged. The aim of this thesis was to create a software solution which could prove helpful in organizing information from a reference and categorize this information correctly ultimately creating an entry in a bib database. This program is not only capable of matching new references with its database but is also able to learn new names which can help with the overall accuracy of the program over time. In conclusion this program serves the purpose it was created for. In the future it can however become a part of a more advanced solution. With appropriate resources the programs usefulness could grow significantly. With the necessary dictionary files the program could be expanded to work with other languages. Also the program might be expanded in the future to support data types that are currently not supported like keywords.

## ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and project KEGA 008TUKE-4/2013: Microlearning environment for education of information security specialists. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] Liddy, Elizabeth D. "Natural language processing." (2001).
- [2] Jewel, Mike. "Paracite developers." Paracite. N.p., n.d. Web. 12 Dec. 2013. <<http://paracite.eprints.org/developers>>.
- [3] Garofalakis, Minos N., Rajeev Rastogi, and Kyuseok Shim. "SPIRIT: Sequential pattern mining with regular expression constraints." VLDB. Vol. 99. 1999.
- [4] Jain, S. et al (1999): *Systems That Learn*, 2nd ed. Cambridge, MA: MIT Press.
- [5] E Hetzner (2008). A simple method for citation metadata extraction using hidden markov models. Proceedings of the 8th ACM/IEEE-CS joint conference on Digital Libraries
- [6] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) *Foundations of Machine Learning*, The MIT Press ISBN 9780262018258
- [7] Davis, Randall. "Knowledge-based systems." *Science* 231.4741 (1986): 957-963.
- [8] Jain, S. et al (1999): *Systems That Learn*, 2nd ed. Cambridge, MA: MIT Press.
- [9] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) *Foundations of Machine Learning*, The MIT Press ISBN 9780262018258
- [10] Michael, Tommi Jaakkola Satinder P. Singh, and I. Jordan. "Reinforcement learning algorithm for partially observable Markov decision problems." *Advances in Neural Information Processing Systems*: 77 (1995): 345.
- [11] Pattern matching of partially ordered events in information management systems / Liberios Vokorokos, Renáta Vokorokosová - 2008. - 1 elektronický optický disk (CD-ROM). In: INES 2008 : 12th International Conference on Intelligent Engineering Systems : Proceedings : February 25-29, 2008, Miami, Florida. - S. l. : IEEE, 2008 P. 169-172. - ISBN 978-1-4244-2083-4

# Odhad Času 3D Modelovania s Využitím Petriho Sietí

<sup>1</sup>František HROZEK, <sup>2</sup>Erik TURANSKÝ

<sup>1,2</sup> Katedra Počítačov a Informatiky, Fakulta Elektrotechniky a Informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

<sup>1</sup>frantisek.hrozek@tuke.sk, <sup>2</sup>erik.turansky@student.tuke.sk

**Abstrakt** — Tento článok sa venuje problematike odhadu času 3D modelovania veľkého množstva budov. Odhad tohto času je nielen základnou časťou prípravy 3D modelovania, ale aj parametrom, ktorý často rozhoduje o tom, či sa dané 3D modelovanie vykoná, alebo nie. V tomto článku je predstavené riešenie, ktoré umožňuje odhad tohto času na základe simulácie s využitím Petriho sietí a nástroja CPN Tools. Článok taktiež predstavuje nadstavbu vytvoreného riešenia, ktorá umožňuje zobrazenie nasimulovaných dát pomocou grafov ako aj porovnanie výsledkov viacerých simulácií.

**Kľúčové slová** — 3D modelovanie, Petriho siete, simulácia, odhad času

## I. ÚVOD

Proces tvorby 3D modelu pozostáva z troch základných krokov [2]: zber a príprava podkladov potrebných na vytvorenie 3D modelu (prípravná fáza), tvorba 3D modelu (modelovacia fáza) a kontrola vytvorených 3D modelov (verifikačná fáza). V súčasnosti existuje mnoho riešení, ktoré zjednodušujú tento proces. Na zber dát je možné použiť satelitné a panoramatické snímky, laserové merače vzdialeností, GPS dáta a iné. 3D model je možné vytvoriť pomocou veľkého množstva aplikácií, ktoré umožňujú vytvorenie tohto modelu v priebehu niekoľkých minút (napr. SketchUp [10], Blender [5]). Na vytvorenie 3D modelu je taktiež možné použiť 3D skener (napr. Leica ScanStation 2 [8], NextEngine 2020i [9]), ktorý dokáže súčasne zozbierať dáta (napr. textúru) a aj vytvoriť 3D model (mračno bodov – point cloud). Prípadne je možné skombinovať 3D skenovanie a 3D modelovacie aplikácie dokopy [1]. Kontrola vytvorených 3D modelov sa vykonáva v 3D modelovacích aplikáciách, v softvéroch dodávaných k 3D skenerom alebo v softvéroch určených pre zariadenia, ktoré budú tieto dáta používať (napr. ZEdit Pro [5] pre 3D tlačiarne od spoločnosti 3D Systems). Môže sa zdať, po vymenovaní všetkých týchto riešení, že proces tvorby 3D modelu je už dostatočne zvládnutý a automatizovaný. Opak je však pravdou. Samotný proces obsahuje ešte mnoho problémov, ktoré je potrebné riešiť. Jedným z nich je odhad celkového času modelovania, ktorý je potrebný v prípravnej fáze. Na základe tohto času je možnú určiť aj cenu tvorby 3D modelu, čo je jedným z hlavných parametrov, podľa ktorého sa rozhodne, či sa dané modelovanie vykoná, alebo nie.

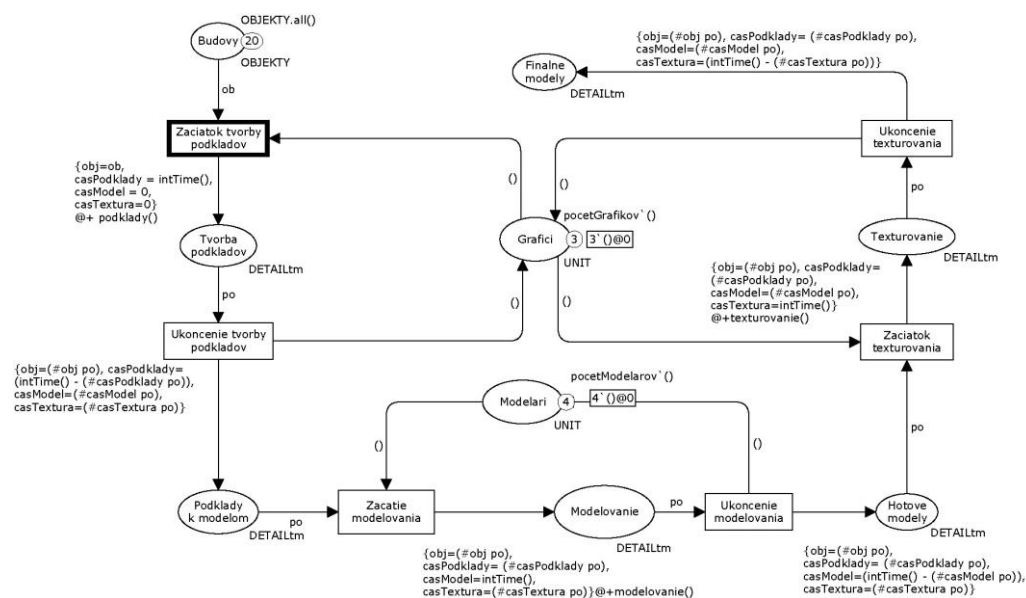
Tento článok ponúka riešenie pre odhad času 3D modelovania budov, ktoré bolo navrhnuté na domácom pracovisku autorov. Toto riešenie využíva simuláciu procesu 3D modelovania na základe identifikovaných parametrov a Petriho sietí. Pre toto riešenie bolo vyvinutá nadstavba, naprogramovaná v jazyku JAVA, ktorá ponúka jednoduchšie nastavovanie jednotlivých parametrov a vizuálnu reprezentáciu a porovnanie nasimulovaných dát.

## II. ODHAD ČASU MODELOVANIA VEĽKÉHO MNOŽSTVA BUDOV

Ako už bolo skôr spomenuté, proces tvorby 3D modelov pozostáva z viacerých krokov, pričom čas ich trvania sa mení podľa druhu modelu ako aj od skúseností pracovníkov. To spôsobuje, že odhad celkového času modelovania je komplikovaný. Ak sa jedná o odhad času modelovania veľkého množstva modelov budov, tento odhad sa ešte viac komplikuje (napr. jednotlivé modely sa môžu veľmi líšiť svojou zložitosťou). Riešením je vytvorenie simulácie, ktorá by na základe zadaných parametrov, odhadla tento čas. V súčasnosti existuje viacero nástrojov, ktoré je možné využiť. Pre vytvorenie tejto simulácie boli použité (časované) farbené Petriho siete (Coloured Petri Nets – CPN) [3][4], ktoré spolu s nástrojom CPN Tools [5] ponúkajú široké možnosti na simuláciu a následnú extrakciu dát.

CPN sú reprezentované orientovaným diparitným grafom (Obr. 1), ktorý pozostáva z miest (kruhy) a prechodov (obdĺžniky). Miesta obsahujú hodnoty, nazývané tokeny a tieto tokeny definujú stav alebo značenie siete. Značenie siete môže byť zmenené odpálením prechodu. Odpálenie prechodu  $t$  nastane, ak je dostatok tokenov v jeho pre-miestach (t.j. miesta z ktorých ide hrana do  $t$ ). Odpálenie odstráni tokeny z pre-miest prechodu  $t$  a vloží tokeny do jeho post-miest (t.j. miesta, do ktorých ide hrana z  $t$ ). Počet a množstvo odobratých a pridaných tokenov je definovaný výrazom zapísaným pri hrane. Detailný popis správania sa CPN modelov je možné nájsť v [1] alebo [5].

Predstavené riešenie pozostáva z množiny CPN simulačných modelov, ktoré popisujú jednotlivé možnosti tvorby 3D modelov bodov. Ukážka jedného z vytvorených modelov (jeden z jednoduchších), ktoré poskytuje navrhnuté riešenie, je znázornená na nasledujúcom obrázku (Obr. 1). Na simulovanie rozdielnosti jednotlivých budov, ako aj ďalších parametrov, ktoré sa využívajú pri týchto simulačných modeloch, bolo použité Gausovo rozdelenie.



Obr. 1 Ukážka jednoduchého časovaného CPN simulačného modelu vytvoreného pomocou CPN Tools

Simulácia pomocou tohto modelu je riadená parametrami znázornenými v nasledujúcej tabuľke (Tabuľka 1).

Tabuľka 1  
 Parametre vzorového CPN simulačného modelu

Názov parametru	Popis parametru
pocetGrafikov	počet grafikov priradených pre tvorbu podkladov a textúrovanie
pocetModelarov	počet modelárov priradených pre tvorbu modelov
poc_objektov	počet objektov, ktoré sa budú modelovať
polygony	predpokladaný počet polygónov objektov
var_poly	odchýlka pre generovanie počtu polygónov objektu Gausovým rozdelením
rych_mod	predpokladaná rýchlosť modelovania (polygóny na hodinu)
var_rych_mod	odchýlka pre generovanie rýchlostí modelovania Gausovým rozdelením (polygóny na hodinu)
cas_jeden_obj	predpokladaná čas tvorby podkladov pre jeden objekt (minúty)
var_cas_jeden_obj	odchýlka pre generovanie času tvorby podkladu Gausovou funkciou (minúty)
textura	predpokladaná rýchlosť textúrovania (minúty)
var_text	odchýlka pre generovanie rýchlosti textúrovania Gausovou funkciou (minúty)

Tvorba 3D modelu, popísaná pomocou tohto CPN simulačného modelu, je rozdelená do troch fáz: tvorba podkladov, modelovanie a textúrovanie. Fáza verifikácie nebola do tohto modelu zahrnutá. Prvá fáza začína odpálením prechodu *Zaciatok tvorby podkladov* a končí odpálením prechodu *Ukoncenie tvorby podkladov*. Nasledovnou fázou je tvorba modelov (prechody *Zacatie modelovania* a *Ukoncenie modelovania*). Poslednou fázou je textúrovanie (prechody



Zaciatok texturovania a Ukoncenie texturovania). Ako je možné vidieť na obrázku (Obr. 1), fáza tvorba podkladov a textúrovanie majú spoločné miesto (*Grafici*). Týmto je dosiahnutá simulácia možnosti, kedy jeden tým grafikov pracuje na oboch fázach súčasne. Časy jednotlivých fáz sú generované pomocou funkcií *podklady*, *modelovanie* a *textúrovanie*, využívajúcich Gausové rozdelenie.

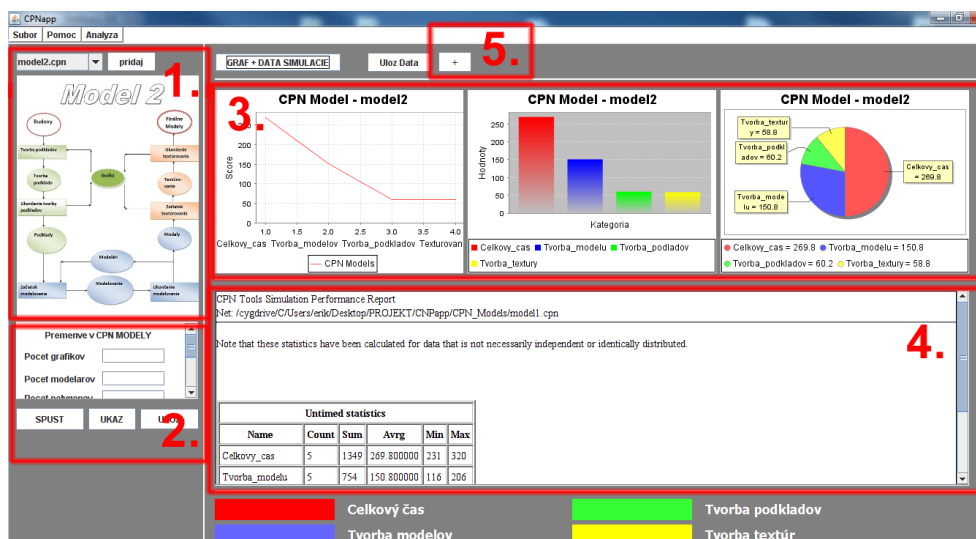
Ukážka výstupu simulácie s využitím tohto CPN modelu je znázornená v nasledujúcej tabuľke (Tabuľka 2). Pre túto simuláciu boli použité nasledovné hodnoty: *pocetModelarov* = 4, *pocetGrafikov* = 3, *poc\_objektov* = 20, *polygony* = 5000, *var\_poly* = 2000, *rych\_mod* = 2500, *var\_rych\_mod* = 1000, *cas\_jeden\_obj* = 5, *var\_cas\_jeden\_obj* = 3, *textura* = 60, *var\_text* = 30.

Tabuľka 2  
 Výstup simulácie

Názov	Počet	Suma	Priemer	Min	Max
Celkovy_cas	20	4934	246.700000	208	280
Tvorba_modelu	20	2517	125.850000	83	163
Tvorba_podkladov	20	1218	60.900000	55	74
Tvorba_textury	20	1199	59.950000	51	69

### III. GRAFICKÉ ROZHRRANIE

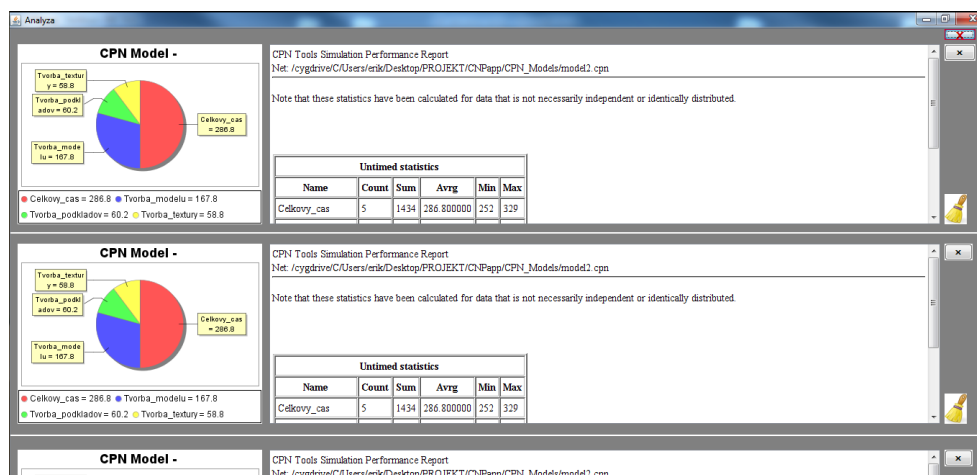
Aj keď sú dáta získané simuláciou zapísané v prehľadnej tabuľke (viď Tabuľka 2), nemusia byť výsledky simulácie hneď viditeľné. Taktiež výsledky viacerých rôznych simulácii nie je možné takto prehľadne zobraziť. Riešením týchto problémov je grafická reprezentácia dát pomocou rôznych druhov grafov, ako aj zobrazenie výsledkov viacerých simulácii vedľa seba. Ďalším problémom pri používaní predstaveného riešenia je skutočnosť, že niektorí používatelia môžu mať problém s nastavovaním parametrov simulácie priamo v CPN Tools. Ako odpoveď na tieto problémy bola vytvorená aplikácia, ktorá tvorí nadstavbu nad existujúce riešenie. Aplikácia bola naprogramovaná v jazyku JAVA a vytvára pre používateľa rozhranie pre prácu s CPN Tools a zobrazovanie nasimulovaných dát. Postup práce s touto nadstavbou je nasledovný. V prvom kroku si používateľ vyberie jednu z vopred vytvorených CPN sietí (Obr. 2 – 1.), prípadne si ich vie pomocou vytvoreného rozhrania upraviť. Následne môže nastaviť simulačné parametre pre vybraný CPN model (Obr. 2 – 2.). Ak je so zvoleným nastavením spokojný, stačí stlačiť tlačidlo *SPUSTIŤ*. Prebehne simulácia v CPN Tools, ktoré aplikácia automaticky spustí, a získané dáta sa zobrazia vo forme grafov (Obr. 2 – 3.). Taktiež sa zobrazí tabuľka, ktorá je štandardným výstupom simulácie v CPN Tools (Obr. 2 – 4.).



Obr. 2 Zobrazenie nasimulovaných dát vo vytvorenej aplikácii

Ak je používateľ spokojný s výsledkami, ale chce vyskúšať aj iné nastavenia parametrov pre danú simuláciu, vytvorená aplikácia mu to umožňuje. Postup je nasledovný. Po odsimulovaní je potrebné uložiť simuláciu na následné porovnanie pomocou tlačidla + (Obr. 2 – 5.). Po uložení môže používateľ spustiť ďalšiu simuláciu, so zmenenými parametrami, a uložiť jej výstupy na následovné porovnanie. Tento krok je možné zopakovať pokiaľ nie je zásobník na porovnávanie

plný (12 simulácií). Uložené simulácie si môže prezrieť po stlačení tlačidla *ANALÝZA*, ktoré sa nachádza v hlavnom menu aplikácie. Následne sa spustí okno analýzy, v ktorom sú pod sebou uložené všetky simulácie, odoslané na porovnanie (Obr. 3). Ak používateľ nie je spokojný s niektorou simuláciou môže ju odstrániť pomocou tlačidla *x*, prípadne môže daný záznam odstrániť a na jeho miesto vložiť ďalšiu simuláciu (pomocou nástroja „metlička“).



Obr. 3 Rozhranie pre porovnanie viacerých simulácií naraz

#### IV. ZÁVER

Odhad času modelovania je základom pre každý projekt, v ktorom je potrebné vytvoriť virtuálnu reprezentáciu vybraných objektov. Avšak tento odhad nie je triviálna záležitosť, keďže doňho vstupuje mnoho parametrov. V tomto článku bolo predstavené riešenie, vytvorené na domácom pracovisku autorov, využívajúce Petriho siete a CPN Tools. V článku bola taktiež predstavená jeho nadstavba, vytvorená v jazyku JAVA, ktorá vznikla ako reakcia na niektoré požiadavky používateľov (grafická vizualizácia dát a porovnanie viacerých simulácií naraz). Výhodou predstaveného riešenia je detailná simulácia pomocou časovaných farbených Petriho sietí (timed CPN) ako aj prehľadne zobrazenie a porovnanie nasimulovaných dát.

V budúcnosti plánujeme rozšíriť predstavené riešenie o možnosť využitia aj iných rozdelení (napr. exponenciálne alebo rovnomerné), ktoré môžu byť použité pri popise niektorých častí procesu tvorby 3D modelovania.

#### POĎAKOVANIE

Táto publikácia vznikla vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: **Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií**, kód ITMS: **26220220182**, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja (80%) a táto práca bola vytvorená realizáciou projektu: **Centrum informačných a komunikačných technológií pre znalostné systémy** (kód ITMS projektu: **26220120020**) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja (20%).

#### REFERENCIE

- [1] F. Hrozek, "3D interfaces of systems," in *Information Sciences and Technologies Bulletin of the ACM Slovakia*, 2013, vol. 5, no. 2, pp. 17-24.
- [2] F. Hrozek, B. Sobota, Cs. Szabo, "Digital preservation of historical buildings using virtual reality Technologies," in *Central European Journal of Computer Science*.2012, vol. 2, no. 3, pp. 272-282.
- [3] K. Jensen, L.M. Kristensen, *Coloured Petri Nets: Modelling and Validation of Concurrent Systems*. Springer, Heidelberg, 2009
- [4] K. Jensen, L.M. Kristensen, L. Wells, "Coloured Petri Nets and CPN Tools for Modelling and Validation of Concurrent Systems," in *International Journal on Software Tools for Technology Transfer*, 2007, vol. 9, no. 3-4, pp. 213-254.
- [5] 3D Systems ZEdit Pro, <http://www.zcorp.com/en/Products/3D-Software/ZEdit/spage.aspx>
- [6] Blender homepage, <http://www.blender.org>
- [7] CPN Tools homepage, <http://cpntools.org>
- [8] Leica ScanStation2 homepage, [http://hds.leica-geosystems.com/en/Leica-ScanStation-2\\_62189.htm](http://hds.leica-geosystems.com/en/Leica-ScanStation-2_62189.htm)
- [9] NextEngine 2020i homepage, <http://www.nextengine.com>
- [10] SketchUp homepage, <http://www.sketchup.com>

# HMM syntéza reči slovenského jazyka s využitím adaptácie modelov

<sup>1</sup>Martin SULÍR, <sup>2</sup>Stanislav ŠOLTÝS, <sup>3</sup>Jozef JUHÁR

<sup>1,2,3</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>martin.sulir@tuke.sk, <sup>2</sup>stanislav.soltys@student.tuke.sk, <sup>3</sup>jozef.juhar@tuke.sk

**Abstrakt** — Syntézu reči možno zaradiť medzi nosné oblasti rečových technológií, a to najmä z dôvodu, že schopnosť systému vytvárať ľubovoľný rečový prejav je nevyhnutná pre čo najrýchlejšiu a najprirodzenejšiu komunikáciu s užívateľom. Predkladaný článok popisuje prvé experimenty s adaptáciou skrytých Markovových modelov pre prípad syntézy reči slovenského jazyka. Pri realizácii experimentov bol použitý existujúci systém syntézy reči v slovenskom jazyku vyvinutý v laboratóriu rečových technológií v telekomunikáciách, ktorý využíva systém Festival a súbor nástrojov HTS. Cieľom bolo získať nové slovenské hlasy s pomerne malými rečovými databázami s využitím technológie adaptácie vopred nadobudnutých modelov a na to určených techník. V článku je postupne opísaný návrh a nahrávanie nových adaptačných databáz na základe foneticky vyváženého textu, adaptácia nových hlasov s využitím modelov priemerného hlasu, ako aj evaluácia novozískaných hlasov pomocou subjektívnych testov.

**KLúčové slová** — adaptácia hlasov, akustická databáza, Festival, HMM syntéza, HTS

## I. ÚVOD

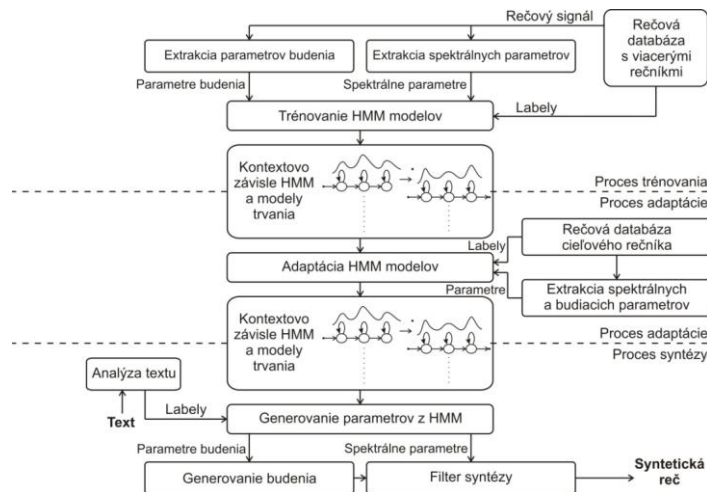
Syntéza reči predstavuje v dnešnej dobe bežnú súčasť života. Stretávame sa s ňou každodenne, napríklad pri cestovaní, či už vlakom, kde sa určitá forma syntetického hlasu využíva pre staničné oznamy, alebo aj pri cestovaní autom, ak využívame satelitnú navigáciu, ktorá disponuje takýmto hlasom. Keďže výpočtová technika je v dnešnej dobe neoddeliteľnou a v podstate aj bežnou súčasťou života, syntézu reči už v súčasnosti predstavujú najmä počítačové systémy. Ich základnou úlohou je uľahčenie života, či už ľuďom s rôznym telesným hendikepom, ako napríklad nevidiacim, alebo aj úplne obyčajným ľuďom, ktorí túto technológiu využívajú najmä na zrýchlenie prenosu a získania požadovaných informácií. Výskum v oblasti rečovej syntézy speje k bodu, kedy budú môcť byť takéto systémy implementované masovo v rôznych oblastiach života bez toho, aby bolo ich použitie limitované z hľadiska ich komplexnosti a celkovej kvality ponúkaného výstupu. Teda výskum je v súčasnosti orientovaný najmä na zlepšovanie kvality z hľadiska zrozumiteľnosti a prirodzenosti prejavu takýchto systémov, ako aj na unifikáciu celého procesu vytvárania nových hlasov.

Tento článok je zameraný na problematiku tvorby nových slovenských hlasov pomocou štatistickej parametrickej syntézy založenej na HMM (*Hidden Markov Models*). Tvorba nových hlasov sa však realizuje pomocou techník adaptácie existujúcich HMM modelov priemerného hlasu pomocou parametrov získaných z adaptačnej databázy. Adaptácia umožňuje použiť veľmi malé akustické databázy pre získanie nového hlasu, teda hlasu s novými parametrami, ako napríklad výškou tónu, intonáciou a pod., pričom parametre nového hlasu zodpovedajú rečníkovi, ktorý nahrával vstupnú adaptačnú databázu.

Článok obsahuje v prvej časti popis syntézy reči na báze HMM s využitím adaptácie modelov. Ďalej je opísaný návrh adaptačných databáz s cieľom maximalizovať ich fonetické pokrytie pomocou predpripraveného foneticky bohatého textu. Táto časť ďalej popisuje nahrávanie, spracovávanie a anotáciu nadobudnutých akustických dát do podoby ucelených akustických databáz. Ďalej je opísaná implementácia systému a adaptácia nových hlasov. V poslednej kapitole sú uvedené výsledky subjektívnych hodnotení nadobudnutých nových hlasov, ktoré boli získané pomocou webového rozhrania a testov na to určených – menovite DMOS (*Degradation Mean Opinion Score*), MOS (*Mean Opinion Score*) a SUS (*Semantically Unpredictable Sentences*) testov.

## II. HMM SYNTÉZA REČI S VYUŽITÍM ADAPTÁCIE MODELOV

Technika adaptácie bola pôvodne vyvinutá pre systémy rozpoznávania reči, kde slúži na adaptovanie všeobecných akustických modelov na špecifického rečníka, tému alebo prostredie, čo má za následok zvýšenie presnosti rozpoznávania. Táto technika našla svoje uplatnenie aj v systémoch pre generovanie reči, kde je jej základnou úlohou vytváranie nových hlasov pomocou malého množstva rečových dát [1]. Obr. 1 zobrazuje blokovú schému HMM syntézy s adaptáciou na rečníka.



Obr. 1 Bloková schéma syntézy reči na báze HMM s adaptáciou rečníka

Syntéza reči na báze HMM modelov s adaptáciou na rečníka (niekedy označovaná aj ako systém pre syntézu reči s využitím modelov priemerného hlasu, angl. AVSS - *Average-Voice-based Synthesis System*) vychádza z pôvodného systému, ktorý bol rozšírení o proces adaptácie. V procese tréningu sa natrénujú HMM modely priemerného hlasu, ktorý vznikol pomocou rozsiahlej rečovej databázy obsahujúcej akustické dáta od viacerých rečníkov. Proces adaptácie hlasov následne pomocou príslušných techník vykonáva modifikáciu modelov na základe menšej rečovej databázy obsahujúcej akustické dáta cieľového rečníka, pričom výstupom tejto úlohy sú modifikované akustické modely zodpovedajúce cieľovému rečníkovi. Finálna syntéza reči prebieha v tomto prípade totožne ako pri prístupe bez využitia adaptácie.

Dve hlavné techniky, ktoré sa využívajú pre adaptáciu na rečníka v prípade systému AVSS, sú metóda maximálnej aposteriornej pravdepodobnosti (MAP - *Maximum A-posteriori Probability*) [2] a metóda maximálnej vierohodnej lineárnej regresie (MLLR - *Maximum Likelihood Linear Regression*) [3].

Využitie adaptácie z pohľadu tvorby nových hlasov robí štatistickú parametrickú syntézu veľmi atraktívnou z hľadiska výskumu a efektívneho využitia takéhoto systému v praktických aplikáciách. Pomocou techniky adaptácie rečníkov je možné s využitím malého množstva dát (stovky viet cieľového rečníka) pomerne jednoducho meniť výsledný hlas, či už z emočného hľadiska, alebo z hľadiska zmeny rečových parametrov.

Prístup s využitím adaptácie modelov ukázal veľký potenciál v tvorbe nových TTS systémov s využitím minimálnych prostriedkov, pričom kvalita dosahovaná na výstupe je v oboch prípadoch adekvátne použiteľná metóde.

## III. SLOVENSKÁ HMM SYNTÉZA REČI S VYUŽITÍM ADAPTÁCIE MODELOV

### A. Návrh adaptačných databáz

V prípade rečovej syntézy môžeme proces vytvorenia nových adaptačných akustických databáz rozdeliť do štyroch nevyhnutných krokov, a to na predspracovanie textu, fonetický výber, nahrávanie a spracovávanie nahrávok. Výsledkom procesu je nadobudnutie akustických dát vo forme nahrávok izolovaných viet.

Základnou úlohou pri tvorbe databáz bol výber optimálnych viet z veľkého textového korpusu [4]. Vybrané vety museli byť foneticky vyvážené, a teda v čo najväčšej miere pokrývať jednotlivé elementy jazyka – v tomto prípade difóny. Proces tvorby spočíval vo filtrácii vstupného textového korpusu, ktorý obsahoval približne 10 miliónov viet [5]. Filtrácia pozostávala z odstránenia viet nevhodných na nahrávanie (príliš dlhých a príliš krátkych viet;

viet, ktoré obsahovali skratky, skomoleniny alebo cudzie slová).

Filtrovaný korpus následne vstupoval do algoritmu selekcie optimálnych viet založenom na *Greedy* selekcii výsledkom čoho bol textový korpus, ktorý obsahoval 330 foneticky vyvážených viet. Tieto vety boli následne nahrávané v profesionálnom nahrávacom štúdiu jedným mužským a jedným ženským rečníkom.

Ďalej boli nahrávky spracované a označené príslušnými značkami – *labelmi*, ktoré definovali hranice jednotlivých elementov v nahrávkach, ako napríklad foném, slabík, slov a pod [6]. Výstupom celého procesu boli dve anotované rečové databázy, jedna s mužským rečníkom a jedna so ženským rečníkom, ktorých vlastnosti sú zobrazené v Tab.1.

Tabuľka 1  
 Parametre databáz pre modely priemerného hlasu

	Mužský hlas	Ženský hlas
<b>Počet viet</b>	330	330
<b>Dĺžka trvania</b>	38 min.	49 min.

### B. Popis databáz pre modely priemerného hlasu

Akustické databázy, ktoré boli potrebné na natréovanie modelov priemerného hlasu vznikli z akustických dát, ktoré sa používajú pre systém rozpoznávania reči v slovenskom jazyku. Tieto akustické dáta boli algoritmicky rozdelené na izolované vety a na jednotlivých rečníkov, ktorý tieto vety nahrávali. Korpus obsahuje vety siedmich ženských a siedmich mužských rečníkov, pričom každý z nich nahovoril viac ako 450 viet. Tieto vety aj s ich textovým prepisom boli vybrané z korpusu pre automatické rozpoznávanie reči, pričom úseky ticha na začiatku a konci každej vety boli orezané na približne 200 ms. Potom bolo potrebné do týchto dát vniesť ortografické čiarky, keďže čiarky v akustických dátach boli čítané. V Tab. 2 sú zobrazené parametre jednotlivých rečových korpusov.

Tabuľka 2  
 Parametre databáz pre modely priemerného hlasu

<b>Databáza mužských rečníkov</b>							
Rečník	1	2	3	4	5	6	7
<b>Počet viet</b>	517	631	543	767	657	531	571
<b>Dĺžka trvania</b>	89 min.	87 min.	102 min.	140 min.	89 min.	107 min.	86 min.
<b>Databáza ženských rečníkov</b>							
Rečník	1	2	3	4	5	6	7
<b>Počet viet</b>	810	545	535	520	570	525	469
<b>Dĺžka trvania</b>	138 min.	97 min.	90 min.	105 min.	112 min.	95 min.	79 min.
<b>Celkový počet viet: 8191</b>				<b>Celková dĺžka trvania: 23 hodín a 36 minút</b>			

### C. Tréovanie modelov a ich adaptácia

Časť tréovania modelov je možné rozdeliť do štyroch základných krokov, a to na extrakciu parametrov, tréovanie modelov, kontextové zhľukovanie pre získanie modelov priemerného hlasu a adaptáciu modelov.

Extrakcia parametrov sa uskutočnila z rečového signálu, ktorý bol vzorkovaný frekvenciou 48 kHz. Extrahovalo sa 34 mel-kepstrálnych koeficientov a logaritmické hodnoty základného tónu logF0 s príslušnými delta a delta-delta koeficientmi.

Parametre modelov boli určené z príslušných rečových dát pomocou metódy využívajúcej kritérium maximálnej vierohodnosti. Použité boli päť stavové ľavo-pravé MSD-HSMM (*Multi-Space probability Distribution-Hidden Semi Markov Model*) modely. V prvom kroku boli natréované robustné modely pre definované monofóny, z ktorých sa následne odvodili kontextovo-závislé modely a nasledoval znovu proces tréovania. Pre čo najrobustnejšiu estimáciu parametrov modelov bolo použité kontextové zhľukovanie založené na kritériu MDL

(*Minimum Description Length*). Rozhodovacie stromy boli v tomto prípade osobitne definované pre spektrálnu (kepstrum) a budiacu ( $\log F0$ ) časť, ako aj pre časť definujúcu dĺžku trvania modelov.

Adaptácia pomocou novovytvorených databáz sa uskutočnila pomocou metódy CMLLR (*Constrained Maximum Likelihood Linear Regression*) v kombinácii s metódou MAP (*Maximum A Posteriori*). V procese adaptácie sa estimovali viacnásobné lineárne transformačné funkcie pomocou vstupných adaptačných databáz mužského a následne aj ženského hlasu. Vzhľadom k tomu, že nie je možné estimovať transformačné funkcie pre všetky zo zhlukov kontextovo-závislých modelov využíva sa tzv. zdieľanie transformačných funkcií pre skupiny súvisiacich modelov.

#### D. Proces syntézy reči

Po adaptovaní HMM modelov nasledoval proces syntézy reči, keď do bloku generovania parametrov vstupovali natrénované HMM modely a *labels*, ktoré vznikli analýzou vstupného textu pomocou nástroja *Festival* a NLP modulu slovenského jazyka [7]. V bloku generovania parametrov sa vygenerovali parametre budenia a spektrálne parametre, a použitím filtra syntézy vznikla syntetická reč. V prípade tohto experimentu bol použitý filter MLSA (*Mel Log Spectrum Approximation*).

### IV. POPIS EXPERIMENTOV

Hodnotenie kvality novovzniknutých hlasov sa realizovalo pomocou subjektívnych posluchovej testov, ktoré spočívali vo vypočítaní nahrávok posudzujúcimi osobami a vyplnením dotazníka. Ten bol vytvorený formou HTML stránky, ktorá obsahovala tri základné testy (DMOS, MOS a SUS). Táto HTML stránka vznikla ako slovenská modifikácia základnej sady nástrojov používanej pri hodnotení *Blizzard Challenge* [8], ktorá sa používa vo svete pre hodnotenie kvality syntézy reči.

Dotazník vo forme webovej stránky sa skladal z troch sekcií pre mužský hlas a troch pre ženský, pričom bola hodnotená prirodzenosť syntetického hlasu testom MOS, jeho podobnosť s pôvodným hlasom pomocou testu DMOS a tiež zrozumiteľnosť pomocou testu SUS.

Vety pre jednotlivé testy boli selektované rovnakým spôsobom pomocou *Greedy* algoritmu ako pri fonetickom výbere textu. Takýto výber bol realizovaný z dôvodu, aby vety neboli vyberané podľa subjektívneho rozhodnutia vhodnosti, ale náhodne. Potom boli syntetizované pomocou systému *Festival* použitím ženského a mužského novovytvoreného hlasu. Jednotlivé syntetizované nahrávky boli rozdelené do dvadsiaticich setov, pričom každý obsahoval 60 viet (10 pre hodnotenie podobnosti mužského hlasu a 10 pre ženský hlas, 10 nahrávok pre hodnotenie prirodzenosti mužského hlasu a 10 pre ženský a nakoniec 10 viet pre test SUS pre oba hlasy).

V prvých dvoch sekciách (pre mužský aj ženský hlas) hodnotili poslucháči podobnosť syntetického hlasu s pôvodným hlasom. Prebiehalo to spôsobom, že účastníci testu mali k dispozícii štyri referenčné nahrávky a k nim desať syntetizovaných viet, ktoré boli pre každého poslucháča iné. Tie sa potom porovnávali s referenčnými nahrávkami a na základe subjektívneho dojmu volili jednu z piatich hodnotiacich možností v škále od 1 do 5, kde 1 znamenala, že daný hlas znel ako úplne iná osoba, a 5, že znel ako úplne rovnaká osoba.

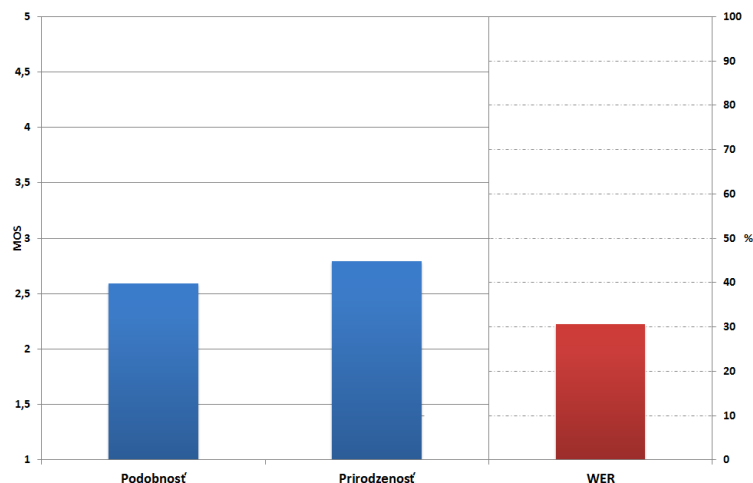
Druhým hodnoteným parametrom bola prirodzenosť syntetizovaných nahrávok, a teda aj samotnej syntézy reči. Pri tomto teste už neboli použité referenčné nahrávky, ale iba syntetizované vety, ktorých bolo znovu desať. Hodnotilo sa rovnakým spôsobom, výberom zo škály od 1 do 5, kde 1 znamenala, že novovytvorený hlas znel úplne neprirodzene, a 5, že znel úplne prirodzene.

Ako posledná sa hodnotila zrozumiteľnosť novovzniknutých hlasov pomocou testu SUS. Jeho cieľom bolo zistiť, či slová, ktoré boli prehraté, boli alebo neboli zrozumiteľné. Každý poslucháč mal poskytnutých desať odlišných viet, ktoré boli syntakticky a gramaticky správne, avšak sémanticky nesprávne, teda nedávali žiaden zmysel. Práve použitie takýchto viet zaručovalo, že účastníci nebudú na základe kontextu dopĺňať automaticky slová, ktorým neporozumeli. Úlohou účastníkov bolo napísať v správnom poradí všetky slová, ktorým porozumeli pri prehrávaní syntetizovaných nahrávok. Na základe vypísaných slov bol vypočítaný pomer chybných slov WER (*Word Error Rate*) ako podiel počtu zle rozoznaných slov ku počtu všetkých [9].

## V. VÝSLEDKY PRVÝCH EXPERIMENTOV

### A. Výsledky hodnotenia mužského hlasu

Subjektívneho hodnotenia kvality syntetickej reči mužského hlasu sa zúčastnilo 30 účastníkov. Tí pomocou mierky DMOS hodnotili podobnosť hlasov syntetickej reči s originálnymi nahrávkami, prostredníctvom mierky MOS prirodzenosť syntetickej reči a pomocou testu SUS jeho prirodzenosť. Získané výsledky boli spriemerované pre nadobudnutie celkového hodnotenia systému a sú zobrazené na nasledujúcom obrázku.

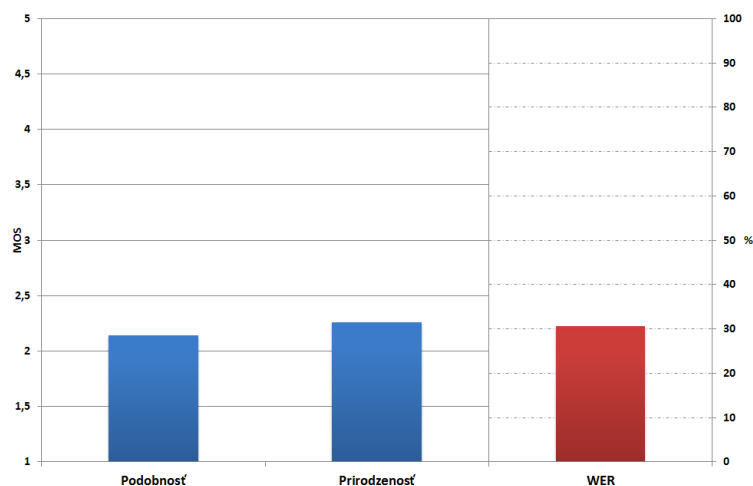


Obr. 2 Získané hodnotenia podobnosti, prirodzenosti a zrozumiteľnosti mužského hlasu

Ako je z hodnotenia zrejmé, podobnosť hlasu syntetickej reči s originálnym nadobudla priemerné hodnotenie 2,59 a prirodzenosť hodnotenie 2,79, pričom hodnotenia sú v rozsahu od 2,5 do 3,5. To predstavuje typický rozsah pre subjektívne hodnotenie syntetickej reči [10]. Zo získaných výsledkov hodnotenia zrozumiteľnosti je zrejmé, že priemerná hodnota chybného rozpoznanie u mužského hlasu bola 30,6%. To znamená, že približne tri slová z desiatich neboli porozumené, resp. boli rozpoznané nesprávne. Ide o pomerne dobrú zrozumiteľnosť systému z hľadiska veľkosti použitej adaptačnej databázy, keďže pomer WER pri obdobných experimentoch v iných krajinách dosahuje v najlepších prípadoch hodnoty v rozmedzí od 15% do 30% [9].

### B. Výsledky hodnotenia ženského hlasu

Rovnakým spôsobom ako pre mužský hlas bolo realizované hodnotenie kvality aj pre ženský hlas. Zúčastnilo sa na tom rovnako 30 poslucháčov, keďže hodnotenie oboch hlasov prebiehalo paralelne v rámci jedného dotazníka. Zhromaždené výsledky sú zobrazené na nasledujúcom obrázku.



Obr. 3 Získané hodnotenia podobnosti, prirodzenosti a zrozumiteľnosti ženského hlasu

Ako môžeme z grafu vidieť, oba testované parametre ženského hlasu nadobudli nižšie hodnotenie, ako to bolo v prípade mužského hlasu (priemerné hodnotenie 2,14 pre podobnosť s originálnym hlasom a 2,26 pre prirodzenosť syntetickej reči). To môže byť zapríčinené inými charakteristikami hlasu a tiež očakávaním, že ženský hlas bude znieť lepšie a prirodzenejšie v porovnaní s mužským.

Výsledky hodnotenia zrozumiteľnosti ukazujú, že priemerná hodnota chybného rozpoznania WER pri ženskom hlase bola rovnaká ako u mužského, teda 30,6 %. V tomto prípade nemala nižšia prirodzenosť ženského hlasu taký dopad na celkovú zrozumiteľnosť, keďže je výsledok testu pri oboch hlasoch rovnaký. Aj keď je výsledné WER trochu vyššie ako priemerne dosahovaný pomer chybovosti vo svete (15% - 30%) [10], je teda možné oba novovytvorené hlasy považovať za zrozumiteľné.

## VI. ZÁVER

V tejto práci sú prezentované prvé experimenty z oblasti využitia adaptácie modelov v štatistickej parametrickej syntéze na báze HMM v slovenskom jazyku. Z nadobudnutých výsledkov je zrejmé, že vytváranie nových hlasov pomocou pomerne malých vstupných adaptačných databáz je možné s dostatočnou kvalitou hlasov na výstupe. Kvalita novovytvorených hlasov je dostačujúca pre úplné porozumenie syntetickej reči, avšak jeho prirodzenosť je na nižšej úrovni, čo je spôsobené najmä použitím filtra pri syntetizovaní požadovaného textu, ktorý spôsobuje nemalú robotickosť hlasu na výstupe. Zvýšenie prirodzenosti je možné použitím pokročilejších metód generovania umelej reči, ako napríklad algoritmu STRAIGHT alebo AHOCODER, čo aj bude predmetom ďalších experimentov.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## POUŽITÁ LITERATÚRA

- [1] T. Masuko, K. Tokuda, T. Kobayashi, S. Imai, "Voice characteristics conversion for HMM-based speech synthesis system", in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP. Munich, Germany, 1997. pp. 1611-1614.
- [2] K. Shinoda, Ch. Lee, "A Structural Bayes Approach to Speaker Adaptation", in IEEE Transactions on Speech and Audio Processing, vol. 9, no. 3., 2001. pp. 276-287.
- [3] J. Yamagishi, T. Kobayashi, N. Yuji, K. Ogata, J. Isogai, "Analysis of Speaker Adaptation Algorithms for HMM-based Speech Synthesis and a Constrained SMAPLR Adaptation Algorithm", in IEEE Transactions on Audio, Speech and Language Processing, vol. 17, no. 1., 2009. pp. 66-83.
- [4] M. Sulír, J. Juhár, „Phonetically Balanced Slovak Speech Corpus for Text-To-Speech Synthesis“, in Proceedings of 13th Scientific Conference of Young Researchers, 2013. pp. 392-394.
- [5] D. Hládek, J. Staš, „Text mining and processing for corpora creation in Slovak language“, in Journal of Computer Science and Control Systems, Vol. 3, No. 1, 2010. pp. 65-68.
- [6] A. Black, K. Lenzo, „FestVox? Building Synthetic Voices“, unpublished. 2007.
- [7] M. Sulír, J. Juhár, "Syntéza reči slovenského jazyka na báze HMM", in Electrical Engineering and Informatics 4 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. Košice, Slovakia, 2013. pp. 39-43.
- [8] A. Black, K. Tokuda, "The Blizzard Challenge – 2005:Evaluating corpus-based speech synthesis on common datasets", in INTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal, 2005. pp. 77-80.
- [9] C. Bennett, A. Syrdal, "Large scale evaluation of corpus-based synthesizers, results and lessons from the 2005 blizzard challenge", in in INTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal, 2005. pp. 85-88.
- [10] X. Huang, A. Acero, H. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm and System Development*, Carnegie Mellon University Pittsburgh: Prentice Hall. 1008 pages. ISBN 0-13-022616-5. 2001.



# Design and Measurements of Microstrip Narrow-band Notch Filter Based on LTCC Substrate Murata LFC

<sup>1</sup>Tibor ROVENSKÝ, <sup>2</sup>Alena PIETRIKOVÁ, <sup>3</sup>Kornel RUMAN

<sup>1,2,3</sup> Department of Technologies in Electronics, Faculty of Electrical Engineering and Informatics, Technical University of Kosice, Slovak Republic

<sup>1</sup>tibor.rovensky@tuke.sk, <sup>2</sup>alena.pietrikova@tuke.sk, <sup>3</sup>kornel.ruman@tuke.sk

**Abstract** — This paper deals with design and comparison between simulations and measurement of microstrip narrow-band notch filter based on LTCC substrates Murata LFC. Design and simulations were provided by Ansoft Designer from ANSYS. Measurements were done by Vector Network Analyser (VNA) up to 10 GHz. The aim of this paper is to compare frequency domain characteristic of fabricated and designed filter based on scattering parameters. Two scattering parameters, return loss (S11) and forward transmission coefficient (S21), were investigated and compared.

**Keywords** — high frequency, microstrip structure, Murata LFC, notch filter

## I. INTRODUCTION

Narrow-band notch filters are used in a lot of applications, for example to minimize interference of other signals at both, receiver and transducer antennas of Ultra Wide Band (UWB) radar system. They are used in many applications like behind the wall detection automotive sensor, medical application etc. [1]. The most important property of UWB radar systems is that they transmit signals across wider frequency in comparison with common radar systems using very low energy level, which is really difficult to detect. Wi-Fi networks are nowadays everywhere and their signal could interference radar systems, therefore, to achieve better accuracy of radar systems, we designed notch filters from 2.4 GHz to 2.5 GHz [2]. Notch filters is designed as a microstrip structure, because in HF application it is not possible to use common electronic elements like resistors, inductors and capacitors.

The biggest challenge of designing narrow-band notch filter is width of stop band, only 100 MHz where Wi-Fi signals are transmitted. If bandwidth of designed filter would be greater it could have as a result losing part of useful signal processed in UWB radar system and overall accuracy of this system could be decreased, therefore, it is necessary to design precise filters with required frequency characteristic.

Microstrip structure of notch filter could be applied by different technologies on various dielectric substrates. Market offers many kinds of substrates but not all of them are suitable for HF applications where stability of dielectric properties (dielectric constant and loss tangent) in large frequency range is needed. Instead of great dielectric properties LTCC substrates are resistant to external influences e.g. vibrations, thermal stress, mechanical stress etc. [3]. These properties were the reason why we have decided to design notch filter on Murata LFC.

## II. NOTCH FILTER

### A. Design of microstrip structure

Notch filter could be realized by different microstrip structures [4], [5]. Final design depends on requirements of filter performance for example, minimal attenuation at stop band, width of stop band, center frequency etc. For narrow-band notch filter requirements were set for minimal attenuation at least -40 dB at stop band, frequency range of stop band is required from 2.4 GHz to 2.5 GHz to minimize only Wi-Fi signals. The best microstrip structure for fulfilling mentioned requirements is assembled from L-resonators which are half wave length long and feed line where signal is transmitted.

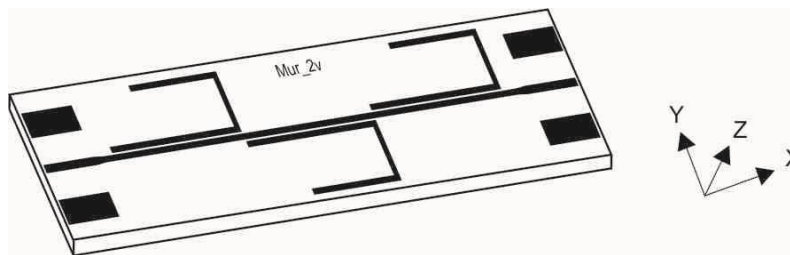


Fig. 1 Microstrip structure of notch filter

Width of microstrip transmission lines is matched to characteristic impedance  $50 \Omega$  to minimize degradation of signal [6]. Pads for connectors which are necessary for measurements are added to the designed structure of notch filter. Microstrip structure of narrow-band notch filter is shown in Fig. 1.

### B. Murata LFC

LTCC substrates are divided into three groups depending on volume of glass. Murata LFC belongs to Glass Ceramic Composite and content from 20 % to 50 % of glass. Amount of glass in LTCC substrate has impact on physical properties and electrical properties depend on ceramic filler. Murata LFC has various areas of applications for example HF modules for communication, power amplifiers, multichip modules and various automotive sensors. Electrical properties of dielectric substrate Murata LFC are shown in Tab. 1 [7].

Table 1  
 PROPERTIES OF MURATA LFC [7]

Property	Symbol	Murata LFC
Dielectric Constant	$\epsilon_r$	7.55
Dissipation Factor Tan	$\Delta$	0.0045

### C. Fabricated notch filter

Notch filter was fabricated by thick film technology. After fabrication process SMA male connectors were soldered to each filter. Soldering process [8], [9] could have impact on microstrip structure performance. Fabricated microstrip notch filter is shown in Fig. 2.



Fig. 2 Fabricated microstrip notch filter (top view)

To reduce costs and minimize possibility of bending dielectric substrate during sintering process full covered ground plane was substituted by mesh ground. During design of mesh ground we paid attention to reduce cost (amount of conductive paste) to minimum without degradation of processed signal. Dimension of fabricated microstrip notch filter without connectors is 56.82 mm x 12.95 mm.

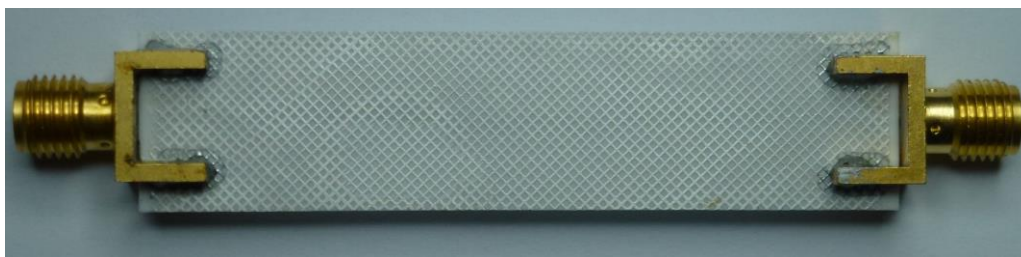


Fig. 3 Fabricated microstrip notch filter (bottom view)

### III. SIMULATIONS AND MEASUREMENTS

Design and simulations were done by Ansoft Designer from ANSYS [10], this software has two options for simulations: linear network analysis and electromagnetic (EM) simulation. EM simulation is more complex and can predict behavior of structures with better accuracy for this reason it was used for simulations. Measurements were done by VNA [11] up to 10 GHz. Two ported VNA has two test cables ended with SMA female connectors. For precision measurements we calibrated VNA before all measurements.

In Fig. 4 comparison between simulated and measured return loss is shown. Divergences between simulation and measured  $S_{11}$  coefficient are caused by change of characteristic impedance. By using thick film technology it is very difficult to fabricate microstrip structure with exactly the same dimensions as it was designed and edges of microstrip are not perfectly sharp which influence characteristic impedance and return losses too. Limitations of thick film technology like resolution, fine printing – shape of edges etc. have to be taken into account by designers.

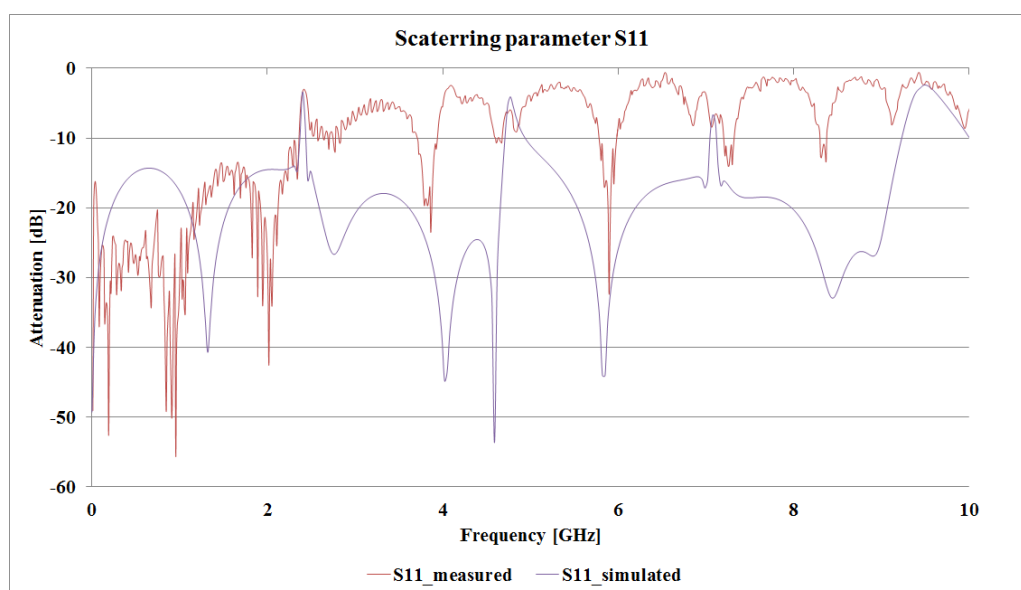


Fig. 4 Simulated and measured  $S_{11}$  parameter

Simulated and measured forward transmission coefficient is shown in Fig. 5. Forward transmission coefficient is influenced by divergences between designed and fabricated device too. Attenuation in stop band is at -37.9 dB, it is less than required but it can be used in UWB applications. Width of stop band is enlarged to 200 MHz from 2.32 GHz to 2.52 GHz. It was caused by enlargement of gaps between feed line and resonators from 150  $\mu\text{m}$  to 170  $\mu\text{m}$ . Width of fabricated feed line and resonators are varied from 0.51  $\mu\text{m}$  to 0.60  $\mu\text{m}$  (designed width of 50  $\Omega$  microstrip line is 0.56  $\mu\text{m}$ ).

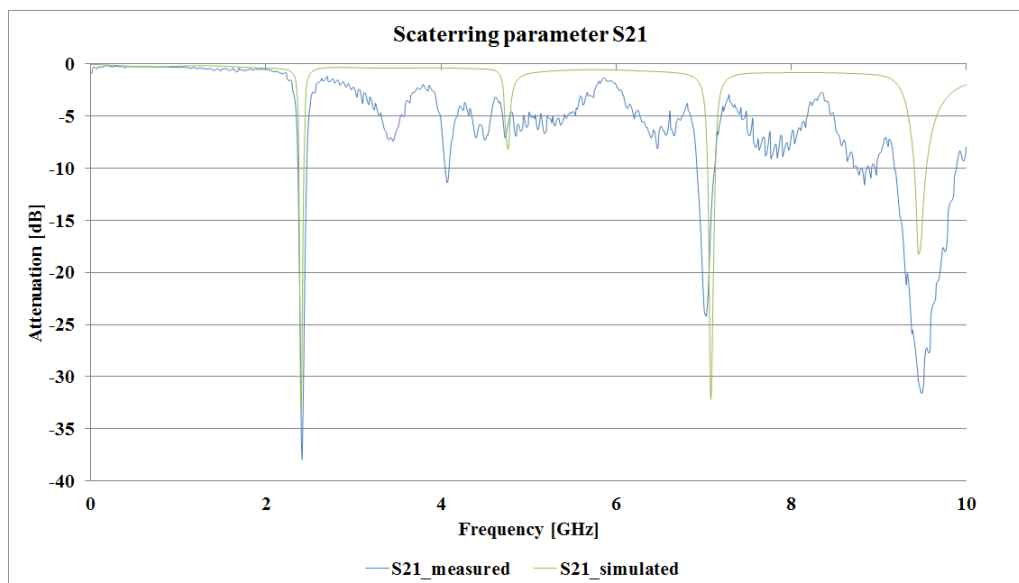


Fig. 5 Simulated and measured S21 parameter

#### IV. CONCLUSION

Microstrip narrow-band notch filter from 2.4 GHz to 2.5 GHz were designed, simulated, fabricated and measured. Return losses and forward transmission coefficient were analyzed and compared with simulation and requirements. Fabricated filter has attenuation only -37.9 dB (required -40 dB) and stop band from 2.32 GHz to 2.52 GHz (required from 2.4 GHz to 2.5 GHz) but in spite these facts it can be used in UWB applications. On the other hand it is necessary to investigate influence of filter with quite different than required performance on UWB radar systems' accuracy.

#### ACKNOWLEDGMENT



Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia/This project is being co-financed by the European Union.

#### REFERENCES

- [1] I. J. Immovreev, and J. D. Taylor, "Future of radars," *Ultra Wideband Systems and Technologies*, 2002, pp. 197-199.
- [2] Wi-Fi / WLAN Channels, Frequencies and Bandwidths, Available on the Internet: 27.6.2014, <http://www.radio-electronics.com/info/wireless/wi-fi/80211-channels-number-frequencies-bandwidth.php>
- [3] R. Alias: "Structural and Dielectric Properties of Glass – Ceramic Substrate with Varied Sintering Temperatures, Sintering Applications", February 2013, ISBN 9789535109747
- [4] Jia-Sheng Hong, "Microstrip Filters for RF/Microwave Applications," John Wiley & Sons, Inc., Hoboken, 2<sup>nd</sup> edition, New Jersey, 2011, 655.
- [5] Jia-Sheng Hong and M. J. Lancaster, "Microstrip Filters for RF/Microwave Applications," John Wiley & Sons, Inc., Hoboken, New York, ISBN 0-471-22161-9, 2001, 471.
- [6] There's Nothing Magic About 50 Ohms, Available on the Internet: 27.6.2014, [http://www.highfrequencyelectronics.com/Archives/Jun07/HFE0607\\_Editorial.pdf](http://www.highfrequencyelectronics.com/Archives/Jun07/HFE0607_Editorial.pdf)
- [7] Low Temperature Co-fired Ceramics (LTCC) Multilayer Module Boards, MURATA LFC, Available on the Internet: 27.6.2014, <http://www.murata.com/products/catalog/pdf/n20e.pdf>
- [8] Rudajevova A., Dusek K., "Study of Undercooling and Recalescence During Solidification of Sn62.5Pb36.5Ag1 and Sn96.5Ag3Cu0.5 Solders in Real Electronic Joints", *Journal of Electronic Materials*, July 2014, Volume 43, Issue 7, pp 2479-2486
- [9] Rudajevova, A., Dusek, K.: "Influence of the thermal history and composition on the melting/solidification process in Sn-Ag-Cu solders", *KOVOVE MATERIALY-METALLIC MATERIALS* Volume: 50 Issue: 5 Pages: 295-300 Published: 2012
- [10] ANSYS DesignerRF, Available on the Internet: 21.6.2013, [http://www.ansoft.com/products/hf/ansoft\\_designer](http://www.ansoft.com/products/hf/ansoft_designer)
- [11] N5221A PNA Microwave Network Analyzer, Available on the Internet: 21.6.2013, <http://www.home.agilent.com/en/pd-2001931-pn-N5221A/pna-microwave-network-analyzer?cc=SK&lc=eng>

# Reasons to Use Formal Methods in Many Areas

<sup>1</sup>Zuzana DUDLÁKOVÁ, <sup>2</sup>Štefan KOREČKO, <sup>3</sup>Branislav SOBOTA

<sup>1, 2, 3</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>zuzana.dudlakova@tuke.sk, <sup>2</sup>stefan.korecko@tuke.sk, <sup>3</sup>branislav.sobota@tuke.sk

**Abstract** — This paper surveys several kinds of formal methods that are used in safety-critical software: Petri nets, VDM, Z notation, B Method, Perfect and Event B. Many questions about their practical usage in the industry have been solved. The paper describes the development and deployment to the practice. Benefits of their using and fields of their application are described. The review can help choose the most appropriate formal method for software or hardware development.

**Keywords** — formal methods, survey, software development, industry

## I. INTRODUCTION

Development of software or hardware system leads to multi-phase and long-term process. All the requirements for the system must be taken into account. They must be combined with functionality, security and velocity. Formal methods have been created to solve or facilitate this entire process.

Today, many computer systems are large and complex. It is expected that their robustness will continue to increase. Failure of systems used in critical areas (medicine, military, energy, transport ...) can have catastrophic consequences in terms of material damage but especially lives.

On this basis, it is very important to ensure the correctness of the systems particularly those that are deployed in critical areas. Ensure their correctness is becoming increasingly difficult through testing, mainly due to their vastness. Testing provides the ability to detect errors incurred during the design and implementation of the system, but it does not guarantee the complete absence of errors [1].

The usage of formal methods is good alternative to testing especially in the design and development process of systems. Formal methods represent a set of strict mathematical processes and tools for writing, transformation and verification of requirements for any system. A correctness of formal specification created in this way is possible to verify and analyse, for example, by proves (B Method, Event B, Perfect ...) or by automatic derivation the system properties from its specification (Petri nets). The usage of formal methods does not guarantee the correctness of the system and it is not customary to apply them to the entire proposal.

## II. OVERVIEW OF THE FORMAL METHODS DEVELOPMENT

So far, researchers have proposed a number of formal methods for the development of computer systems. Examples include: Petri nets [2], VDM [3], Z notation [4], B Method [5], Event B [6] and Perfect language [7].

The design of hardware is one of the first areas where formal methods have found their application [8]. Their use, whether in design control or simulation, contributed to their commercial availability.

Benefits of using formal methods are considerably dependent on their smart applications. Their usage is not automatic. There is no single best way to use formal methods and any the best method. Moreover, formal methods are only part of the problem solution in software development. Their success depends mainly on their integration into a larger process [9].

Reasons to use formal methods may include the following [10]:

- The main problem is the need to solve the system as a whole, from its design, implementation, through testing, up to the introduction into service, not just the

software proposal.

- The combination of modern technology with formal methods can help in the development of safety-critical software.
- Their usage in practice leads to an improvement the quality and time and it reduces costs.

#### A. Development of formal methods

Since 1990, many questions about practical use of formal methods in industrial processes have been solved. Several significant explorations were made where the problems related to verification were identified. They were the subject of further research in an effort to resolve them. The following table (Table 1) shows a historical overview of the development of formal methods and their application in the industry since 1990.

Table 1  
 Overview in the development of formal methods

Year	Situation
1990	Hall [11] expressed 7 myths about formal methods as a defence of their use in the industry.
1990	Wing [12] explained the basic terms and principles of formal methods for novices.
1992	Thomas [13] presented proves of formal methods cost-efficiency in industrial applications.
1993	Austin and Parkin [14] did an exploration of industrial use of formal methods in order to determine the reasons for their relatively low acceptance in the industry.
1993	Craigen, Gerhart and Ralston [15], [16] did an exploration of the application of formal methods in practice, which consists of the twelve case studies. All of them had positive conclusions.
1993	Rushby [17] wrote a technical report for NASA about usage of formal methods in development of critical aircraft systems.
1996	There were many different opinions even though the optimistic conclusions of some researches. Hinchey and Bowen [18] wanted that the standards of formal methods, their tools and education have been accepted in the industry. But Glass [19] saw a large gap between academics who take formal methods as a necessity and practitioners who see formal methods as irrelevant.
1997	Campos and Harrison [20] made an overview in progression of formal verification usage in interactive systems.
1999	Bloomfield and Craigen [21] did an extensive exploration involving the evaluation of research programs, significant conferences and industrial application areas. Both of them saw benefits in software proposal of critical systems. They also identified several reasons for rejection of formal techniques. They noted that the jejuneness of theories and tools means that some success applications require heroic level of effort.
2009	Jim Woodcock [22] did report about formal methods usage in 62 projects from around the world in 2008-2009. His results indicate an effectiveness of formal methods usage for time, cost and quality. An improvement through the use of formal methods became better in 35% for time, in 37% in costs and in 92% for quality.
2011	Silva [23] used Petri nets to describe the flow in discrete event models.

#### B. Selected formal methods

Z notation, B Method and VDM (Vienna Development Method) are model-based languages, which usually model a system through its states. They are based on set theory and mathematical logic. This state model is created using mathematical entities such as sets, relations, sequences and functions. Operations of the system define the change of its states [24].

Other newer formal methods are Perfect and Event B. Perfect language is based on the concept of classes like object-oriented software. Event B was developed as an evolution of B Method.

Petri nets represent the only formal language that enables an automated derivation of invariant properties of system. They are successful use in modelling various applications, for

example communication systems, fault-tolerant systems and also parallel systems.

An overview of mentioned formal methods with their authors is in the following table (Table 2). The figure (Fig. 1) shows the evolution of aforementioned formal methods on the timeline together with their associated tools and languages.

Table 2  
 Overview of selected formal methods

Formal method	Year	Authors
Petri nets	1962	Carl Adam Petri
VDM	70s of the 20 <sup>th</sup> century	IBM institute in Vienna
Z notation	1977	J. R. Abrial, S. Schuman, B. Meyer
B Method	1985 – 1988	J. R. Abrial
Perfect	2003	David Crocker
Event B	2005	J. R. Abrial

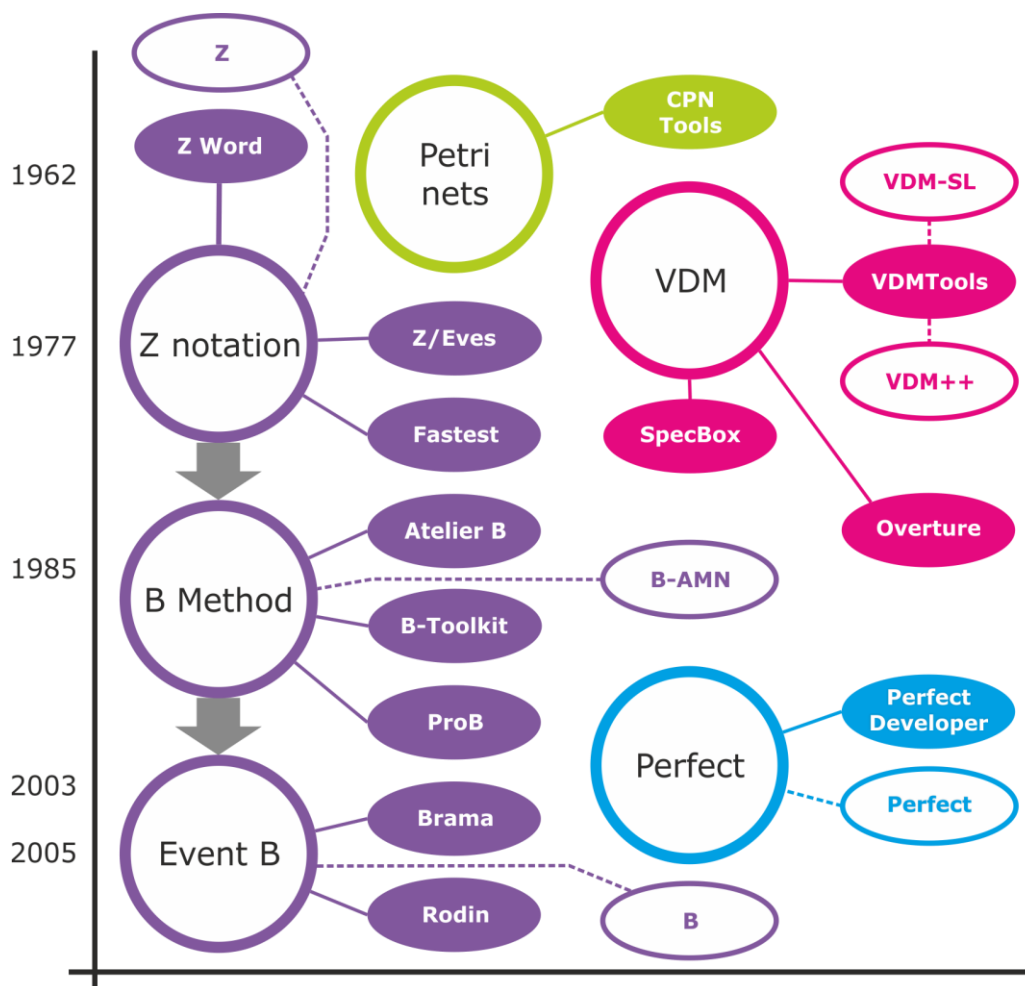


Fig. 1 Overview of the development of selected formal methods

### III. FIELDS OF APPLICATION OF FORMAL METHODS

Formal methods are nowadays used in many industries. They have the most benefit in the construction of safety-critical systems (aircraft systems, train systems, nuclear power plants, medical systems, storage and usage of weapons), but also in the microprocessor design. Formal methods have also found their application in systems for controlling ATM cards, in virtual reality systems and in computer games.

#### *A. Aircraft systems*

One of the primary examples of the use of formal methods in practice is the project SIFT (Software Implemented Fault Tolerance) [25]. It is probably one of the most important American experiments in safety-critical sector. SIFT is a very reliable aircraft control computer. It was ordered by NASA in 70s with very strict safety requirements. Formal verification consisted of two phases: Boyer-Moore theory prover and Shostak theory prover

Airbus uses SCADA (Safety Critical Application Development Environment) technology [26] for development of on-board computers. Technology is on the market from 2003 and it provides automatic generation of safe and quality code in languages ADA 83/95, SPARK and DO-178B. It enables a precise formal proposal of system that is deterministic and without ambiguity. Big advantages are the availability in rapid changes of specification or code modification, simulation and easy to learn. The costs of the development halved with the introduction of this technology.

#### *B. Train systems*

System SACEM (Systeme d'Aide a la Conduite, a l'Exploitation et a la Maintenance) [27] was developed in 80s for the Paris metro. Its technique was powerless against the design errors and standard validation through testing was considered as unsuccessful. This was the reason for using a formal approach, where the source code enriched with pre/post conditions and it was subsequently verified. Specification of software was rewritten in formal language, which was manually controlled. Also, this technique was very expensive to develop, so in 1998 the control system Météor [28] got into operation for the metro line No. 14 in Paris. This application was the first which used the B Method. It proved to be very successful because all errors were detected during the development. A validation process was carried out very efficient.

In 2007, the company ClearSy developed an application in B Method for Siemens Transportation Systems. It ensured an automated shuttle to airport.

The company Alstom Transport Information Solutions uses B Method and Event B in software development of train systems.

#### *C. Nuclear power plants*

Formal methods have found their application in the area of nuclear power plants [29]. Companies such AREVA NP (France), Schneider Electric (France), BARC (India), Rolls-Royce Civil Nuclear (France, United Kingdom), KAERI (Korea) and others use SCADA technology built on formal methods for instrumentation and control application as well as for security applications in the nuclear industry. Currently, there are 37 operating nuclear reactors (France, Armenia, Belgium, China, Czech Republic, Lithuania). Their control parts have been developed through SCADE technology.

#### *D. Medical systems*

Some medical instruments have vitally critical function and they require a high degree of reliability. Formal specifications have been used in devices for monitoring the vital functions of patients in intensive care and surgery rooms. They also found an application in improving the electrocardiogram (ECG) [30]. The accuracy increased by using formal methods in development process. And several problems and ambiguities were found in formal specification of the product.

#### *E. Storage of weapons*

Holding of large amounts of ammunition is managed by ACS (Ammunition Control Software) [31] system, which has undergone extensive validation. It is important that all of the safety rules are implemented by system because technicians are forced to rely on the system output. Safety requirements are formulated in Orange Book. The formal method VDM was used for their formal specification. Moreover, this research has demonstrated the success of enlargement of formal methods into areas with planning and prediction.

#### *F. Microprocessors*

VIPER (Verifiable Integrated Processor for Enhanced Reliability) [32] microprocessor was specially developed for safety-critical applications. HOL (Higher Order Logic) was used for verification of critical parts and this microprocessor was first, which was subjected to formal verification. It satisfies the requirements for quality and simplicity. Formal verification was used



in the design of various types of processors (ARM processors, current microprocessors, VLIW microprocessors). At present, formal methods are very widespread and they are used in the design of various parts of the hardware.

#### G. ATM cards

In 90s, an electronic cash system Mondex (smart cards) [33] was developed. It was suitable for low values of cash and it had no transaction costs. It was very important that the card was safe and the electronic money could not be counterfeited. Extensive proves were generated by Z notation. They revealed several errors and the proposal of system had to be amended.

#### H. Virtual reality

Validation and verification techniques have been identified as suitable mechanisms for determining whether the software satisfies the needs of user, for verification whether the software works properly. However, existing verification techniques do not support friendly visualization. Also, validation methods with friendly visualization mechanisms do not allow verifying the accuracy of system. In 2009, the method for the validation and verification of software systems was presented through the integration of formal methods and virtual reality [34].

#### I. Computer games

In recent years, the area of computer games is exposed to considerable interest with regard to the processing of data, modelling, creating and design. Petri nets are often used to modelling gaming behaviour to control the problems arising in the design of games. Modelling using Petri nets [35] make a way for designers to detect problems before the game is actually developed. This can be especially important in the simulation of multi-players games or heterogeneous games, where it is difficult to predict the dynamic behaviour of the game.

### IV. CONCLUSION

There are several reasons in practical usage of formal methods. B Method as a formal method is the most widespread especially in the design of critical parts of the transport systems. Petri nets have the greatest application in the areas where the description of the system dynamics is needed. They are optionally needed in modelling parallel processes or in visualisation of simulation of solution through virtual reality.

Formal methods find an application in the industry, especially in the design of safety-critical software in order to avoid the errors. Many errors are often detected only after the introduction into service. Their late detection could lead to irreversible damage or loss of the human life. Thanks to new technologies and the development of virtual-reality systems there are many various areas for potential application of formal methods (medicine, military, virtual-reality applications, computer games, simulations), but especially in area of the transport systems that are arranged between safety-critical applications.

### ACKNOWLEDGMENT

This work has been supported by KEGA grant project No. 050TUKE-42012: "Application of Virtual Reality Technologies in Teaching Formal Methods". Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

### REFERENCES

- [1] DIJKSTRA, E. W.: Why Correctness must be a Mathematical Concern. In: R. S. Boyers, J. S. Moore, editors, The Correctness Problem in Computer Science, Academic Press, 1981.
- [2] PETRI, C. A.: Introduction to General Net Theory. In: Proceedings of the Advanced Course on General Net Theory of Processes and Systems, Hamburg, 1979, Springer-Verlag, 1980, SS. 1-19.
- [3] LARSEN, P. G., MUKHERJEE, P., PLAT, N., VERHOEF, M., FITZGERALD, J.: Validated Designs for Object-oriented Systems. In: Springer Verlag, 2005. ISBN: 1852338814
- [4] SPIVEY, J. M.: The Z Notation: A Reference Manual. In: Prentice Hall International, Oxford, UK, 1992.
- [5] ABRIAL, J. R.: The B book: Assigning programs to Meaning. In: Cambridge University Press, New York, 1996.
- [6] MÉTAYER, C., ABRIAL, J. R., VOISIN, L.: Event-B Language. In: Rodin, 31 May, 2005.
- [7] CROCKER, D.: Teaching Formal methods with Perfect Developer.
- [8] HEITMEYER, C.: On the Need for Practical Formal Methods. In: Proceeding in 5th International Symposium, FTRTFT'98 Lyngby, Denmark, September 14-18, 1998 S. 18-26. ISBN 978-3-540-65003-4
- [9] HALL, A.: Realising the Benefits of Formal Methods. In: Proceeding 7th International Conference on Formal Engineering Methods, ICFEM 2005, Manchester, UK, November 1-4, 2005, S. 1-4. ISBN 978-3-540-29797-0
- [10] BOWEN, J., STAVRIDOU, V.: Safety-critical systems, formal methods and standards. In: Software Engineering Journal, July, 1993, S. 189-209.

- [11] HALL, A.: Seven Myths of Formal Methods. In: IEEE Softw. 7, September, 1990. S. 11-19
- [12] WING, J. M.: A specifier's introduction to formal methods. In: IEEE Comput. 23, 1990. S.8-24
- [13] THOMAS, M.: The industrial use of formal methods. Microproc. Microsyst. 17, 1992. S.31-36
- [14] AUSTIN, S., PARKIN, G.: Formal methods: A survey. In: Tech. rep., National Physical Laboratory, Teddington, Middlesex, UK, 1993.
- [15] CRAIGEN, D., GERHART, S., RALSTON, T.: An International Survey of Industrial Applications of Formal Methods. Vol. 1 Purpose, Approach, Analysis and Conclusions. In: U.S. Department of Commerce, Technology Administration, National Institute of Standards and Technology, Computer Systems Laboratory, Gaithersburg, MD, 1993.
- [16] CRAIGEN, D., GERHART, S., RALSTON, T.: An International Survey of Industrial Applications of Formal Methods. Vol. 2 Case Studies. In: U.S. Department of Commerce, Technology Administration, National Institute of Standards and Technology, Computer Systems Laboratory, Gaithersburg, MD, 1993.
- [17] RUSHBY, J.: Formal methods and the certification of critical systems. In: Computer Science Laboratory, Menlo Park, CA, December 1993.
- [18] HINCHEY, M. G., BOWEN, J. P.: To formalize or not to formalize? In: IEEE Comput. 29, April 1996. S.18-19
- [19] GLASS, R. L.: Formal methods are a surrogate for a more serious software concern. In: IEEE Comput. 29, April 1996.
- [20] CAMPOS, J. C., HARRISON, M. D.: Formally Verifying Interactive Systems: A Review. In: Department of Computer Science, University of York, Heslington, Springer, 1997, S. 109-124.
- [21] BLOOMFIELD, R., CRAIGEN, D.: Formal methods diffusion: Past lessons and future prospects. In: Adelard, Coborn House, London, UK, December 1999.
- [22] WOODCOCK, J., LARSEN, P. G., BICARREGUI, J., FITZGERALD, J.: Formal Methods: Practice and Experience. In: ACM Computing Surveys, Vol. 41, No. 4, Article 19, October 2009.
- [23] SILVA, M., JÚLVEZ, J., MAHULEA, C., VÁZQUEZ, C. R.: On fluidization of discrete event models: observation and control of continuous Petri nets. In: Springer Science, September 2011.
- [24] KAUR, A., GULATI, S., SINGH, S.: Analysis of Three Formal Methods-Z, B and VDM. In: International Journal of Engineering Research & Technology, Vol. 1, Issue 4, 2012. ISSN: 2278-0181
- [25] WENSLEY, J. et al.: SIFT: Design and Analysis of a Fault-tolerant Computer for Aircraft Control. In: Proceeding in IEEE, 1978, S. 1240-1254.
- [26] BERRY, G.: Synchronous design and verification of critical embedded systems using SCADE and Esterel. In: Proceedings of the Formal Methods for Industrial Critical Systems. Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg Germany, 2008.
- [27] HENNEBERT, C., GUIHO, G.: SACEM: a fault tolerant system for train speed control. In: 23rd International Conference on Fault-Tolerant Computing, Toulouse, France, 1993.
- [28] BEHM, P., BENOIT, P., FAIVRE, A., MEYNADIER, J. M.: Météor: a Successful Application of B in a Large Project. In: FM'99, Toulouse, France, 1999.
- [29] HILL, J. V.: The development of high reliability software – RRGAs' experience for safety critical systems. In: Second IEE/BCS Conference Software Engineering '88, July 1988, S. 169-172.
- [30] LADEAU, B. R., FREEMAN, C.: Using formal specification for product development. In: Hewlett-Packard Journal, 1991, S. 62-66.
- [31] MUKHERJEE, P., STAVRIDOU, V.: The formal specification of safety requirements for the storage of explosives. In: Technical Report DITC 185/91, National Physical Laboratory, Teddington, Middlesex, UK, August 1991.
- [32] CULLYER, W. J., PYGOTT, C. H.: Application of formal methods to the VIPER microprocessor. In: IEE Proc. E., 1987, S. 133-141.
- [33] STEPNEY, S., COOPER, D., WOODCOCK, J.: An electronic purse: Specification, refinement, and proof. Technical Monograph PRG-126. In: Oxford University Computing Laboratory, 2000.
- [34] ALMEIDA, H., SILVA, L., FERREIRA, G., LOUREIRO, E., PERKUSICH, A.: Validation and Verification of Software Systems Using Virtual Reality and Coloured Petri Nets. In: Federal University of Campina Grande, Brasil, 2009.
- [35] CHAMPAGNAT, R., PRIGENT, A., ESTRAILLIER, P.: Scenario building based on formal methods and adaptive execution. In: Université de La Rochelle, France, 1 February, 2005.

# Kategorizácia textových dokumentov v slovenskom jazyku metódami kontrolovaného učenia

*Daniel ZLACKÝ, Ján STAŠ, Jozef JUHÁR, Anton ČIŽMÁR*

Katedra elektroniky a multimediálnych telekomunikácií,  
Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach,  
Park Komenského 13, 042 00 Košice, Slovenská republika

{daniel.zlacky, jan.stas, jozef.juhar, anton.cizmar}@tuke.sk

**Abstract** — Tento článok sa zaoberá kategorizáciou slovenských textových dokumentov rôznymi metódami kontrolovaného učenia, ako napríklad rozhodovacie stromy, naivný Bayesov klasifikátor, metóda maximálnej entropie alebo Winnov algoritmus. K dispozícii sme mali viac ako 4 700 slovenských dokumentov, ktoré boli zatriedené do 5 kategórii. Pri testovaní jednotlivých metód sme využili techniku krížovej validácie. Najlepšie výsledky klasifikácie z hľadiska f-miery dosiahla metóda maximálnej entropie.

**Keywords** — klasifikácia dokumentov, naivný Bayesov klasifikátor, rozhodovacie stromy, Winnov

## I. ÚVOD

S narastajúcim počtom textových dokumentov, článkov alebo kníh na Internete je problematika kategorizácie týchto údajov stále aktuálna vec. Predchádzajúci výskum sa venoval algoritmom nekontrolovaného učenia. Jednotlivé metódy a techniky zhukovania textových dokumentov boli porovnané a zhodnotené vo viacerých článkoch. Dokázali sme, že vhodným predspracovaním a výberom zhukovacej metódy je možné korpus textových dokumentov rozdeliť do ucelených tried a natrénovať jazykový model adaptovaný na určitú oblasť, ktorý dosiahne lepšie výsledky v systéme automatického rozpoznávania reči [1].

Tento článok však bude venovaný problematike kontrolovaného učenia pri klasifikácii jednotlivých textových dokumentov. K dispozícii máme trénovací korpus, ktorý bol rozdelený do určitých kategórii podľa témy obsiahnutej v daných dokumentoch. Naším cieľom je porovnať rôzne klasifikačné algoritmy a v ďalšom výskume uplatniť dosiahnuté poznatky pri kategorizovaní väčšieho textového korpusu za účelom skvalitnenia jazykového modelu v systéme automatického rozpoznávania reči [2].

## II. KLASIFIKÁCIA DOKUMENTOV

Klasifikácia textových dokumentov predstavuje techniku, ktorá automaticky zatriedi textové dokumenty do jednej alebo viacerých kategórii na základe použitého klasifikačného algoritmu. Problém klasifikácie textových dokumentov môže byť definovaný nasledovne. Predpokladajme, že máme množinu trénovacích textových dokumentov  $D = \{d_1, \dots, d_n\}$ , kde každému dokumentu je priradená trieda z množiny  $k$  rôznych diskretných hodnôt indexovaných  $\{1 \dots k\}$ . Trénovacia množina je použitá na vytvorenie klasifikačného modelu, ktorý každému neznámemu dokumentu z testovacej množiny priradí prislúchajúcu triedu. V nasledujúcich riadkoch si rozoberieme niekoľko základných algoritmov, ktoré sme použili pri kategorizácii textových dokumentov.

### A. Naivný Bayesov klasifikátor

Nech  $D$  je vyššie definovaná množina dokumentov. Každý dokument je potom reprezentovaný príznačným vektorom, kde každá položka tohto vektora obsahuje príznačenie konkrétneho slova (term) zo slovníka  $V = \{t_1, \dots, t_m\}$ . Každý dokument je zároveň členom aspoň

jednej triedy z množiny všetkých tried  $C = \{c_1, \dots, c_c\}$ . Neznámy dokument je potom zaradený do triedy na základe výpočtu podmienených pravdepodobností pre každú triedu a výberom triedy s maximálnou aposteriornou pravdepodobnosťou. Pravdepodobnosť, že dokument  $d$  bude zaradený do triedy  $c_j$  (podmienená aposteriorná pravdepodobnosť) je definovaná Bayesovým pravidlom:

$$P(c_j|d) = \frac{P(c_j)P(d|c_j)}{P(d)}$$

$P(c_j)$  predstavuje apriórnu pravdepodobnosť, ktorá je vypočítaná z množiny tréningových dát ako relatívny výskyt dokumentov patriacich do triedy  $c_j$ . Podmienená pravdepodobnosť  $P(d|c_j)$  dokumentu  $d$  vzhľadom k triede  $c_j$  je vypočítaná z pravdepodobnosti výskytu jednotlivých slov ku všetkým slovám obsiahnutých v dokumente  $d$  nasledovne:

$$P(d|c_j) = \frac{|d|!}{\prod_{t \in d} f(t, d)} \prod_{t \in d} P(t|c_j)^{f(t, d)}$$

kde  $|d|$  predstavuje súčet výskytu všetkých slov v dokumente  $d$ ,  $f(t, d)$  udáva počet výskytov slova  $t$  v dokumente  $d$  a  $P(t|c_j)^{f(t, d)}$  vyjadruje pravdepodobnosť jednotlivých slov vzhľadom ku všetkým slovám v danej triede. Dokument je potom zaradený do príslušajúcej triedy výberom triedy s maximálnou aposteriornou pravdepodobnosťou  $P(c_j|d)$  nasledovne [3]:

$$c_{\max} = \arg \max_{c_j \in C} P(c_j|d) = \arg \max_{c_j \in C} P(c_j) \prod_{k=1}^n P(t_k|c_j)$$

### B. Rozhodovacie stromy

Metódy využívajúce rozhodovacie stromy na klasifikáciu textových dokumentov sú založené na hierarchickom rozklade tréningového priestoru na základe určitých vlastností a podmienok, ktoré na výpočet využívajú výskyt alebo absenciu jedného alebo viacerých slov v dokumentoch. Delenie tréningového priestoru v rozhodovacích stromoch je vykonávané rekurzívne, pokiaľ koncový uzol obsahuje určitý minimálny počet objektov, ktoré charakterizujú danú triedu. Na každý testovací dokument sú potom aplikované tieto podmienky a vlastnosti, na základe ktorých sa nájde optimálna cesta v rozhodovacom strome ku koncovému uzlu, ktorý charakterizuje dovtedy neznámy testovací dokument. Aby sa predišlo pretrénovaniu, niektoré koncové uzly môžu byť z rozhodovacieho stromu odstránené pomocou informácie získanej z množiny odložených dát (množina dokumentov, ktoré neboli použité pri tréningu rozhodovacieho stromu). V prípade klasifikácie textu môžeme dokumenty kategorizovať ku koncovým uzlom pomocou výskytu alebo absencie charakteristických slov v dokumente. Ako uzly si v tomto prípade môžeme predstaviť slová alebo frázy, ktoré sú medzi sebou maximálne odlišné [4].

### C. Metóda maximálnej entropie

Metóda maximálnej entropie využíva tréningové dáta k obmedzeniu rozloženia podmienenej pravdepodobnosti. Predpokladajme, že máme  $n$ -rozmerný vektor príznakov, ktorý odpovedá funkcii  $f_i(d, c)$  pre daný dokument  $d$  a triedu  $c$ . Našou úlohou je vytvoriť model, ktorý bude tomuto rozložению príznakov odpovedať. Nech  $D$  je množina všetkých tréningových dokumentov. Hľadaná podmienená pravdepodobnosť  $P(c|d)$ , ktorá priradí dokument  $d$  do triedy  $c$  musí spĺňať nasledujúcu vlastnosť:

$$\frac{1}{|D|} \sum_{d \in D} f_i(d, c(d)) = \sum_d P(d) \sum_c P(c|d) f_i(d, c)$$

kde  $P(d)$  označuje pravdepodobnosť rozloženia dokumentov, ktorú nepoznáme. Na aproximáciu tejto pravdepodobnosti použijeme množinu tréningových dokumentov:

$$\frac{1}{|D|} \sum_{d \in D} f_i(d, c(d)) = \frac{1}{|D|} \sum_{d \in D} \sum_c P(c|d) f_i(d, c)$$

Naledujúcim krokom by bolo vypočítanie príznakových funkcií z tréningových dát, ktoré by boli použité na klasifikáciu. Potom pre každý príznak vypočítať očakávanú hodnotu nad tréningovými údajmi, ktorá bude použitá ako obmedzenie tréningového modelu [5].

### D. Winnow algoritmus

Winnow algoritmus patrí do triedy lineárnych inkrementálnych algoritmov a je veľmi podobný algoritmu perceptronov. Hlavný rozdielom je to, že pri perceptronoch sa jednotlivé

váhy aditívne zvyšovali o určitú konštantu, avšak pri algoritme Winnow sa násobia. Na začiatku sú všetky váhy  $w = \{w_1, \dots, w_n\}$  inicializované na rovnakú kladnú hodnotu. Algoritmus takisto inicializuje aj prahovú hodnotu  $\theta$ . Váhy sa menia len ak je dokument klasifikovaný do nesprávnej triedy. Hodnota o akú sa zmenia dané váhy je závislá od parametra  $\alpha > 1$ . Vhodné prahové hodnoty dosiahneme ak nastavíme prah na hodnotu  $\theta = n/2$  a  $\alpha = 2$ . Podrobnejšie je daný algoritmus a jeho modifikácie opísané v [6].

### III. KORPUS TEXTOVÝCH DOKUMENTOV

Trénovací a testovací korpus slovenských textových dokumentov bol zostavený z novinových článkov, ktoré boli automaticky stiahnuté z Internetu a zatriedené do 5 kategórii. Detailný popis tejto problematiky je uvedený v článku [7]. V Tabuľke 1 je možné vidieť, že daný korpus pozostáva celkovo z 4 744 textových dokumentov, ktoré sú zatriedené do kategórii ako Ekonomika, Šport, Správy, Svet a Zdravotníctvo. Najväčší počet dokumentov je obsiahnutých v kategórii Správy a naopak najmenší v kategórii Zdravotníctvo.

Tabuľka 1  
 Štatistiky korpusu textových dokumentov

Kategória	Počet dokumentov	Počet viet	Počet slov	Počet jedinečných slov
<b>Ekonomika</b>	1 020	20 636	376 035	35 384
<b>Šport</b>	1 037	25 175	328 397	38 823
<b>Správy</b>	1 650	26 770	455 722	59 683
<b>Svet</b>	478	8 284	149 005	19 557
<b>Zdravotníctvo</b>	559	13 020	229 907	14 057
<b>Spolu</b>	4 744	93 885	1 539 066	167 504

Keďže náš korpus pozostával len zo 4 744 dokumentov, rozhodli sme sa pri testovaní jednotlivých klasifikačných techník použiť krížovú validáciu. Korpus bol teda rozdelený na 10 približne rovnakých častí, z ktorých 9 častí bolo stále použitých na trénovanie a 1 časť na testovanie klasifikačnej metódy. Týmto spôsobom sme jednu klasifikačnú metódu 10 krát testovali stále na inej časti korpusu, ktorá nebola použitá v procese trénovania klasifikátora. Získané výsledky predstavujú aritmetický priemer čiastkových výsledkov.

#### A. Evaluácia

Na evaluáciu dosiahnutých výsledkov sme sa rozhodli použiť základné miery, ktoré sa vo väčšine prípadov používajú pri hodnotení výsledkov kategorizácie textových dokumentov. Najprv si však rozoberieme, aké rôzne prípady môžu nastať pri klasifikácii dokumentov:

- TP – True Positive – Dokument bol správne zaradený do danej kategórie
- TN – True Negative – Dokument bol správne nezaradený do danej kategórie
- FP – False Positive – Dokument, ktorý mal byť zaradený do danej kategórie bol systémom zaradený do inej kategórie
- FN – False Negative – Dokument, ktorý mal byť zaradený do inej kategórie bol systémom priradený do danej kategórie

Jednotlivé prípady sú ilustrované aj v nasledujúcej Tabuľke 2, označovanej ako konfúzna matica.

Tabuľka 2  
 Konfúzna matica pri binárnej klasifikácii

	Predikovaná kategória	
Správna kategória	TP	FN
	FP	TN

Z danej tabuľky vychádzame aj pri definovaní štyroch základných hodnotiacich mier:

- Presnosť („Accuracy“) – vyjadruje percento správne zaradených dokumentov do danej kategórie

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}$$

- Precíznosť („Precision“) – vyjadruje pomer medzi správne zaradenými dokumentmi do danej kategórie ku súčtu týchto dokumentov s dokumentmi, ktoré nepatria do

danej kategórie

$$Pcn = \frac{TP}{TP + FP}$$

- Návratnosť („*Recall*“) - vyjadruje pomer medzi správne zaradenými dokumentmi do danej kategórie ku súčtu týchto dokumentov s dokumentmi, ktoré mali byť zaradené do danej kategórie, ale systémom boli priradené do inej

$$Rcl = \frac{TP}{TP + FN}$$

- F-miera („*F-measure*“) – komplexnejšia miera, ktorá vyjadruje kompromis medzi precíznosťou a návratnosťou

$$F\_msr = \frac{2 * Pcn * Rcl}{Pcn + Rcl}$$

#### IV. EXPERIMENTÁLNE VÝSLEDKY

V rámci experimentálneho výskumu kategorizácie textových dokumentov sme sa rozhodli otestovať 6 klasifikačných algoritmov. Dosažené výsledky z hľadiska presnosti a f-miery sú ilustrované v Tabuľke 3. Na tréning a testovanie jednotlivých algoritmov bol použitý nástroj s názvom Mallet [8]. Z tabuľky je možné vidieť, že najlepšie výsledky boli dosažené s klasifikátormi ako Vyvážený Winnow algoritmus, metóda maximálnej entropie a naivný Bayesov klasifikátor s hodnotami f-miery v rozmedzí 0,93 – 0,94. Naopak najhoršie klasifikačné výsledky sme dosiahli pri použití rozhodovacích stromov.

Tabuľka 3  
 Experimentálne výsledky

Metóda	Presnosť [%]	F-miera
<b>Vyvážený Winnow algoritmus</b>	93,40	0,93
<b>Rozhodovacie stromy</b>	55,73	0,51
<b>Metóda maximálnej entropie</b>	94,62	0,94
<b>Naivný Bayesov klasifikátor</b>	93,23	0,93
<b>MC metóda maximálnej entropie</b>	80,39	0,76
<b>Winnow algoritmus</b>	77,93	0,78

#### V. BUDÚCI VÝSKUM

V rámci budúceho výskumu chceme dosažené výsledky využiť pri klasifikovaní rozsiahleho textového korpusu, ktorý sa využíva v systéme automatického rozpoznávania reči [2]. Správnu kategorizáciou tréningových dokumentov môžeme dosiahnuť zníženie chybovosti daného systému a zvýšiť jeho robustnosť.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] D. Zlacký, J. Staš, J. Juhár, A. Čížmár, „Text categorization with Latent Dirichlet Allocation“, In *Proceedings of the 10th International Conference on Advances in Electro-Technologies*, Oradea, Romania, 2014.
- [2] J. Juhár, J. Staš, D. Hládek, „Recent progress in development of language model for Slovak large vocabulary continuous speech recognition“, In *New Technologies: Trends, Innovations and Research*, pp. 261-276, Rijeka, Croatia, 2012.
- [3] C. D. Manning, P. Raghavan, H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 1st Edition, 2008.
- [4] J. R. Quinlan, „*Induction of decision trees*“, In *Machine Learning* 1, pp. 81-106, 1986.
- [5] K. Nigam, J. Lafferty, A. McCallum, „Using maximum entropy for text classification“, In *Workshop on Machine Learning for Information Filtering IJCAI-99*, pp. 61-67, 1999.
- [6] N. Littlestone, „Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm“, In *Machine Learning* 2, pp. 285-319, 1988.
- [7] D. Hládek, J. Staš, J. Juhár, „The Slovak categorized news corpus“, In *Ninth International Conference on Language Resources and Evaluation LREC 2014*, Reykjavik, Iceland, pp. 1705-1708, 2014.
- [8] A. K. McCallum, *MALLET: A Machine Learning for Language Toolkit*, 2002.

# Vzdialené pracovisko pre FPGA obvody Altera

<sup>1</sup>Lukaš VOJTKO, <sup>2</sup>Martin PETRVALSKÝ, <sup>3</sup>Miloš DRUTAROVSKÝ

<sup>1,2,3</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky Technickej univerzity v Košiciach

<sup>1</sup>lukas.vojtko@student.tuke.sk, <sup>2</sup>martin.petrvalsky@tuke.sk, <sup>3</sup>milos.drutarovsky@tuke.sk

**Abstrakt** – V tomto príspevku opisujeme vypracované webové rozhrania, užívateľské príručky a testovacie úlohy vo VHDL jazyku k jednotlivým cvičeniam pre vzdialené testovanie navrhnutých úloh. Na začiatku je vysvetlená štruktúra projektu eDiViDe. Cvičenia sa venujú problematike návrhu konečného automatu, posuvného registra s lineárnymi spätnými väzbami a skutočného generátora náhodných čísel.

**Kľúčové slová** – projekt eDiViDe, vzdialené pracovisko, FPGA obvody Altera, konečný automat, posuvný register s lineárnymi spätnými väzbami, skutočný generátor náhodných čísel

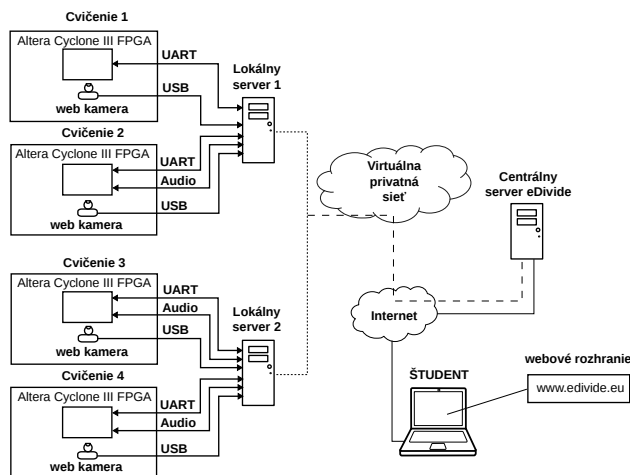
## I. ÚVOD

Návrh digitálneho systému je v dnešnej dobe veľmi rozšírený a napreduje veľkou rýchlosťou. Väčšina zariadení pracuje s digitálnymi dátami a preto vzdelávanie v tejto oblasti je nevyhnutné pre návrhárov elektronických systémov. Rozvojom elektronických číslicových obvodov, ako sú FPGA (programovateľné hradlové polia) čipy, sa rozšírili možnosti návrhov digitálnych obvodov. V týchto čipoch je možné navrhnuť vlastné obvody, periférie a mikroprocesory, ktoré najlepšie vyhovujú požiadavkám navrhovaného systému. Vývojové dosky s čipom FPGA sú drahé a preto je k dispozícii na katedre malý počet takýchto dosiek. Aby sa tieto vývojové dosky využili naplno a boli kedykoľvek k dispozícii pre študentov, katedra sa zapojila do projektu eDiViDe (European Digital Virtual Desing Lab) [1], ktorý spája viacero vysokých škôl s cieľom vytvoriť celoeurópske vzdialené laboratórium digitálneho dizajnu. Laboratórium ponúka rôzne druhy cvičení so sprievodnými úlohami a príkladmi pre vývojové dosky s čipom FPGA od firmy Altera. Našou úlohou bolo navrhnuť webové rozhranie a zostaviť úlohy a príklady pre sadu aplikácií vytvorených na našej katedre a tiež vytvorenie užívateľského manuálu.

## II. PROJEKT EDIVIDE A POUŽITÝ HARDVÉR

### A. Štruktúra projektu eDiViDe

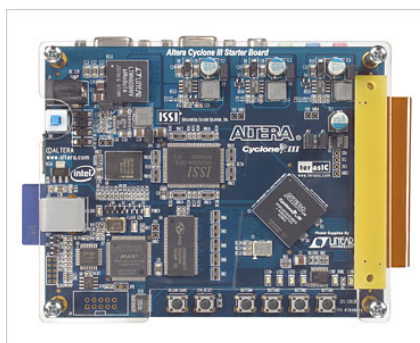
Projekt eDiViDe spája viacero vysokých škôl s cieľom vytvoriť celoeurópske vzdialené pracovisko pre testovanie návrhov na báze obvodov FPGA. Na centrálnom serveri sú umiestnené stránky jednotlivých projektov, ktoré si každá vysoká škola zostaví na základe svojich projektov. Prostredníctvom centrálnej stránky sa študent dostane k jednotlivým projektom. Ku každému projektu je vytvorených niekoľko úloh. Postupným riešením úloh sa študent zoznami s problematikou daného cvičenia. Študent podľa inštrukcií k úlohe zostaví svoj návrh v jazyku VHDL. Jednotlivé návrhy môže študent simulovať a navrhovať na svojom počítači prostredníctvom návrhového prostredia pre cieľové obvody FPGA [8]. Každý študent bude mať svoj prihlasovací účet, prostredníctvom ktorého sa prihlási do svojho konta. Študent svoj VHDL (jazyk pre opis hardvéru) návrh úlohy nahrá na server. Pre nahranie na dosku a otestovanie je potrebné rezervovať čas, kedy bude môcť študent svoj návrh testovať. Na testovanie je vyhradený pevný časový interval (aktuálne 10 minút). Pre každé cvičenie je vytvorená samostatná stránka, ktorá je tvorená virtuálnymi perifériami pre ovládanie každej úlohy. V jednotlivých úlohách môžu byť k dispozícii všetky ovládacie prvky danej stránky alebo len ich časť. Obmedzenia budú vopred k danej úlohe definované a bude uvedený aj postup ako jednotlivé periférie využiť a ako ich pripojiť k vytváranému návrhu. Katedra má aktuálne štyri projekty rozdelené na dve jednoduché a dve zložitejšie cvičenia. Celá štruktúra projektu eDiViDe pre Technickú univerzitu v Košiciach aj s naznačenou komunikáciou je znázornená na Obr. 1.



Obr. 1 Štruktúra a prepojenie technických prostriedkov projektu eDiViDe.

### B. Hardvér lokálnych pracovísk

Jednotlivé vývojové dosky typu Altera NIOS II Embedded Evaluation Kit, Cyclone III Edition [2] pozostávajú z dosky obsahujúcej čip Altera Cyclone III EP3C25F324 FPGA (Obr. 2). Vývojové dosky komunikujú so serverom prostredníctvom UART (Univerzálny asynchrónny prijímač a vysielač) rozhrania. Lokálne servery obsahujú web kameru Logitech, Inc. Webcam C310 a externé audio karty typu C-Media Electronics, Inc. Audio Adapter pre LiveStream. USB Blaster je súčasťou vývojovej dosky, pomocou ktorého sa nakonfiguruje návrh do obvodu FPGA od firmy Altera.



Obr. 2 Hardvérová platforma Nios II Cyclone III edícia.

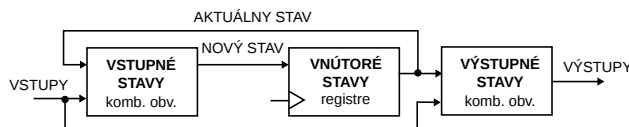
## III. VYTVORENÉ CVIČENIA

Vytvorili sme sadu cvičení, ktoré študent bude riešiť v niekoľkých úlohách. Študent vytvorí návrh obvodu vo VHDL jazyku. Pre každé cvičenie je vytvorené webové rozhranie prispôbené pre jednotlivé cvičenia. Týmto spôsobom sa študent naučí navrhovať obvody na vývojových doskách FPGA v jazyku VHDL.

### A. Cvičenie 1 - Ovládanie 7-segmentového displeja pomocou konečného automatu

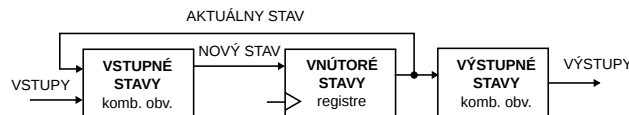
Prvé jednoduché cvičenie študenta zoznami s rôznymi typmi FSM (konečný automat) [7][8]. Pomocou FSM bude študent ovládať virtuálny 7-segmentový displej. Nakoniec by mal byť študent schopný navrhnuť kuchynský budík. Prvý konečný automat, s ktorým sa študent zoznami je typu Mealy. Jeho štruktúra je zobrazená na Obr. 3.





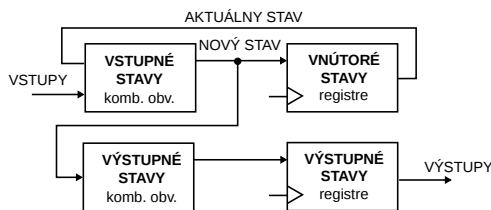
Obr. 3 Konečný automat typu Mealy

Ďalší konečný automat je typu Moore, ktorého štruktúra je zobrazená na Obr. 4.



Obr. 4 Konečný automat typu Moore

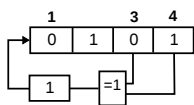
Posledným typom konečného automatu, s ktorým sa študent zoznami, je RNS (registrovanie nasledovných stavov). Hlavnou vlastnosťou konečného automatu typu RNS je, že jeho nový vnútorný stav sa registruje. Výstupné signály sú potom platné skôr, ako pri ostatných typoch konečných automatoch. Štruktúra konečného automatu typu RNS je zobrazená na Obr. 5.



Obr. 5 Konečný automat typu RNS

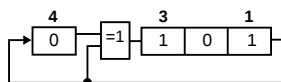
### B. Cvičenie 2 - Generátor šumu mora pomocou LFSR

Druhé jednoduché cvičenie je zamerané na problematiku návrhu obvodu LFSR (posuvného registra s lineárnymi spätnými väzbami)[8]. Študent sa zoznami s rôznymi typmi LFSR a bude ich musieť využiť vo svojich návrhoch. V poslednej úlohe využije svoje skúsenosti pre vytvorenie generátora šumu mora pomocou LFSR. Základným aspektom je pochopiť LFSR, čo je N-bitový posuvný register, ktorý ma  $2^N - 1$  stavov. Nasledujúci stav je lineárnou funkciou predchádzajúceho stavu. Bity ktoré definujú ďalší stav sú jednotlivé bloky posuvného registra zodpovedajúce navrhnutému polynómu. Tým, že je daná počiatočná hodnota registra a že činnosť LFSR je deterministická, stavy produkované LFSR sú úplne určené. LFSR má preto konečný počet možných stavov a nakoniec sa tieto stavy budú cyklicky opakovať. LFSR s dobre zvoleným blokmi posuvného registra môže produkovať sekvenciu bitov, ktoré majú veľmi dlhý cyklus. Fibonacciho LFSR má na základe polynómu privedené výstupy posuvného registra na vstup XOR-u (vylučovací OR) a tie sú postupne do seba vnorené a výstup je privedený na LSB alebo MSB LFSR. Štruktúra LFSR typu Fibonacci je zobrazená na Obr. 6.



Obr. 6 Fibonacciho 4-bitový LFSR.

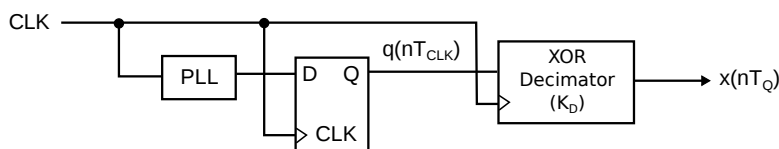
Galoisov LFSR má tiež na základe polynómu privedené bloky posuvného registra na vstupy XOR-ov, ale jednotlivé výstupy sú pripojené k častiam posuvného registra. Štruktúra LFSR typu Galois je zobrazená na Obr. 7.



Obr. 7 Galoisov 4-bitový LFSR.

### C. Cvičenie 3 - TRNG koprocessor pre NIOS Soft Core

V prvom zložitejšom cvičení ma za úlohu študent navrhnuť TRNG (skutočný generátor náhodných čísel) generátor [3] na báze PLL (fázový záves), ktorého štruktúra je znázornená na Obr. 8. PLL obvody sú vygenerované pomocou funkcie MegaWizard v prostredí Altera Quartus II [4]. Základným princípom metódy TRNG je získať náhodnosť extrahovanú z nestability (jitter) hodinového signálu z PLL. Pomocou soft procesora Nios II [6], ktorý je implementovaný na čipe FPGA, sa vyhodnocujú FIPS (Federal Information Processing Standards) testy. Následne sú výsledky posielať na webové rozhranie pre vizualizáciu. Teória a ďalšie informácie sú k dispozícii v [5].



Obr. 8 Princíp spoľahlivej extrakcie náhodnosti z PLL.

## IV. ZÁVER

V príspevku je opísaný projekt eDiViDe a jednotlivé cvičenia, ktoré Katedra elektroniky a multimediálnych telekomunikácií pridala do tohto projektu. Tým, že v dnešnej dobe digitálne systémy sú veľmi rozšírené a FPGA obvody dávajú návrhárom veľké možnosti, je veľmi dobré, že sa katedra zapojila do projektu eDiViDe. Študent má v týchto cvičeniach uľahčenú prácu tým, že sa nezaobera komunikáciou a funkcionalitou použitých modulov v štruktúre projektu eDiViDe. Hlavnou výhodou je fakt, že jednotlivé vývojové dosky budú k dispozícii cez vzdialené pracovisko.

## POĎAKOVANIE

Tento príspevok bol vytvorený v rámci projektu European Digital Virtual DEsign lab (EDIVIDE) #518565-LLP-1-2011-1-BE-ERASMUS-ESMO, ktorý financuje LLP Erasmus.

## LITERATÚRA

- [1] *eDiViDe*. (2013) eDiViDe home webpage. [Online]. Available: <http://www.edivide.eu>.
- [2] Altera Corporation, *Cyclone III Device Handbook*, Altera, v 12.0, August 2012.
- [3] V. Fischer and M. Drutarovsky, "True Random Number Generator Embedded in Reconfigurable Hardware," in B.S. Kaliski, Jr., C.K. Koc, C. Paar (Eds.): "Cryptographic Hardware and Embedded Systems," 4th International Workshop, CHES 2002, Redwood Shores, California, USA, August 13-15, 2002, LNCS 2523, Springer, Berlin, 2003, pp.415-430
- [4] Altera Corporation, "Altera Phase-Locked Loop (AL- TERA PLL)" Megafunction User Guide, Altera, January 2011.
- [5] "Remotely Testable Setup of Soft CPU with Cryptographic TRNG Coprocessor Extension Embedded into Altera FPGA" Dept. of Electronics and Multimedia Communications Technical University of Kosice Kosice, Slovakia
- [6] Altera Corporation. (2013) "Nios II Embedded Evaluation Kit, Cyclone III Edition User Guide", Altera, July 2010. [Online]. Available: <http://www.altera.com/products/devkits/altera/kit-cyc3-embedded.html>.
- [7] V. Fischer, "VHDL Language - Design and synthesis of digital systems". presentation 2010.
- [8] L. Vojtko, "Vzdialené pracovisko pre FPGA obvody Altera". Bakalárska práca, Technická univerzita v Košiciach, Fakulta elektrotechniky a informatikyn, 2014.

# Viackanálový prenos dát vo frekvenčnom pásme SRD

<sup>1</sup>Štefan HEDVIG, <sup>2</sup>Martin PETRVALSKÝ, <sup>3</sup>Miloš DRUTAROVSKÝ

<sup>1,2,3</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovensko

<sup>1</sup>stefan.hedvig@student.tuke.sk, <sup>2</sup>martin.petrvalsky@tuke.sk,  
<sup>3</sup>milos.drutarovsky@tuke.sk

**Abstrakt** — Článok sa zaoberá riešením bezdrôtového komunikačného rozhrania pre sieť UWB radarov. Opisuje návrh hardvéru, vytvorenie komunikačného protokolu, programovanie softvéru pre zadaný procesor a vytvorenie obslužných programov pre počítač. Cieľom bolo vytvorenie fungujúceho riešenia, ktoré je možné ďalej rozširovať. Navrhnutý bol hardvér zariadenia pripravený pre ďalšie rozšírenie, následne sme vytvorili softvér, testovali riešenie a hľadali chyby a nedostatky súčasného riešenia.

**Kľúčové slová** — Ultra-širokopásmové radary UWB, Pásmo zariadení krátkého dosahu SRD, ARM mikroprocesory, bezdrôtová sieť

## I. ÚVOD

Bezdrôtové systémy na prenos dát sú dnes v značnej miere rozšírené a na trhu existuje veľké množstvo hotových riešení. Nakoľko sú však požiadavky kladené na bezdrôtový systém prenosu dát pre senzorovú sieť UWB radarov špecifické, vytvorili sme riešenie navrhnuté priamo na mieru radarového systému [1]. Toto riešenie je navrhnuté ako modulárne, navrhnuté tak aby bolo možné jednoducho ďalej systém prispôbovať požiadavkám. Vytvorenie kompletného systému vyžaduje riešenie niekoľkých samostatných problémov.

V prvej časti článku analyzujeme spôsoby komunikácie a frekvenčné pásmo, v ktorom bude zariadenie komunikovať. Následne sme špecifikovali ako bude potrebné vybudovať hardvér. Pre bezdrôtový komunikačný systém sú najdôležitejšie práve využité moduly pre bezdrôtovú komunikáciu.

Po návrhu a optimalizácii požiadaviek na hardvér je možné realizovať návrh hardvérového riešenia, na ktorom je možné začať vývoj softvéru. Opísali sme návrh a výrobu výsledného zapojenia a jeho využitie pri tvorbe a ladení softvéru. V neskoršej vývojovej fáze softvéru sme pristúpili k stanoveniu definitívnych požiadaviek na hardvér.

Následne sa venujeme opisu funkcií softvéru, opisu jeho tvorby, vlastností, možnosti úprav a spojeniu s ostatnými časťami radarového systému. Pre prenos dát do bezdrôtového komunikačného systému a spracovanie dát na jeho výstupe sme navrhli programové vybavenie pre platformu PC. V závere článku sú zhodnotené dosiahnuté výsledky a je naznačený smer ďalšieho vývoja.

## II. REALIZÁCIA PRENOSU DÁT

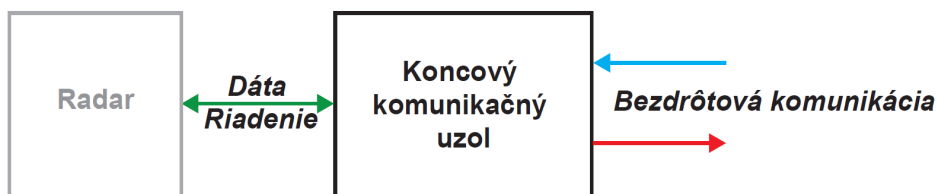
### A. Komunikačné rozhranie pre radarovú sieť

UWB radarová sieť pozostáva z niekoľkých samostatných radarov, ktoré je potrebné prepojiť s hlavnou jednotkou, ktorá spracúva údaje zo senzorov. Radary využívajú pre prenos dát do dátového koncentrátora samostatný kanál pre každý uzol siete. Po tomto kanáli sa prenáša väčšina dátovej komunikácie a je teda potrebné zabezpečiť prenosový kanál pre každý uzol dostupný vždy, keď ho potrebuje. Spätný - potvrdzovací kanál je realizovaný spoločným vysielateľom z dátového koncentrátora. V spätnom smere sa realizuje iba malá časť komunikácie a to potvrdzovanie prijatia paketov a režia spojenia. Je teda možné s výhodou využiť spoločný komunikačný kanál pre všetky 4 radary. Navrhnutý komunikačný systém teda využíva ku komunikácii 5 frekvenčných kanálov vo frekvenčnom pásme SRD [2].

### B. Návrh hardvérovej konfigurácie

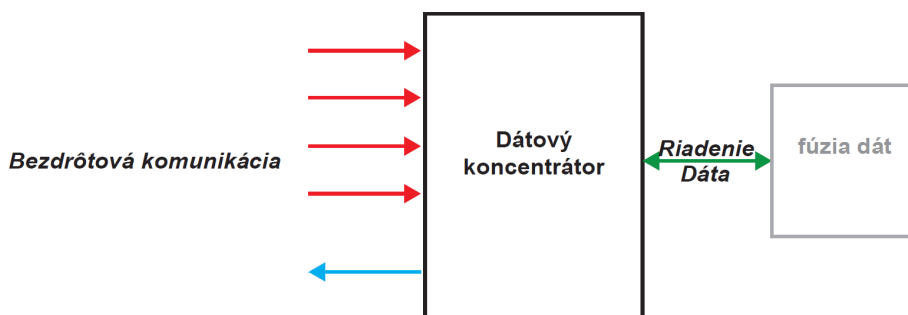
Pre vytvorenie systému bezdrôtovej komunikácie bolo potrebné navrhnuť koncepciu pripojenia rádiových modulov a periférií a v nasledujúcom kroku aj úplnú hardvérovú konfiguráciu dvoch zariadení. V komunikačnom systéme sa nachádzajú dva typy zariadení.

**Komunikačný uzol** sa pripája k senzoru radarovej siete, kde sprostredkúva dátovú komunikáciu. Jeho primárnou úlohou je zabezpečiť prenos dát zo senzora do systému pre spracovanie dát. Komunikuje s jedným zariadením - dátovým koncentrátorom a teda vyžaduje k funkcii jeden prijímací a jeden vysielač rádiový modul. Bloková schéma komunikačného uzla je zobrazená na Obr. 1.



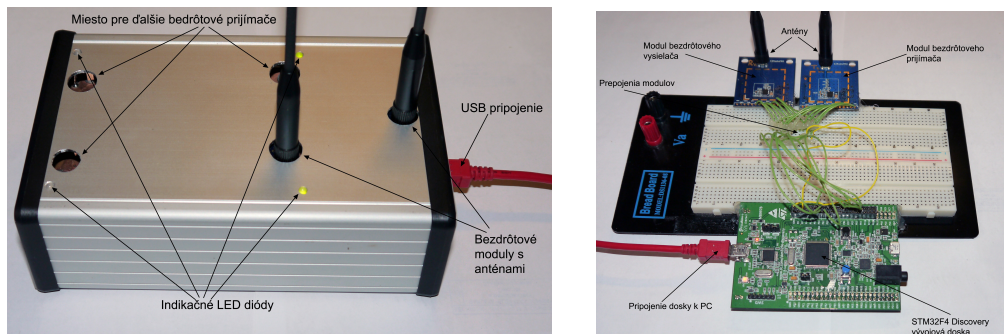
Obr.1 Bloková schéma komunikačného uzla

Druhým prvkom je **dátový koncentrátor**, ktorého úlohou je príjem dát zo 4 komunikačných uzlov. Pre zabezpečenie požiadavky spojenia so 4 uzlami sme navrhli osadenie koncentrátora 4 prijímačmi a jedným spoločným vysielačom. Prijímače musia vždy byť pripravené na začatie komunikácie na dohodnutom kanáli, je teda potrebné mať funkčný prijímač pre každý komunikačný uzol. Dátový koncentrátor vysiela však iba občasné malé množstva dát - potvrdenie prijatia údajov a komunikáciu spojenú s nadviazaním, kontrolou a uzatvorením spojenia. Z tohto dôvodu bol zvolený jeden spoločný vysielač, čím sa dosahuje úspora komunikačných kanálov. Celkové zapojenie zobrazuje bloková schéma na Obr. 2.



Obr.2 Bloková schéma dátového koncentrátora

Hardvér bol navrhnutý a postavený na procesoroch s jadrom ARM [3] od STM. Bola využitá vývojová doska STM32F4Discovery [4]. Výsledný hardvér je zobrazený na Obr. 3.

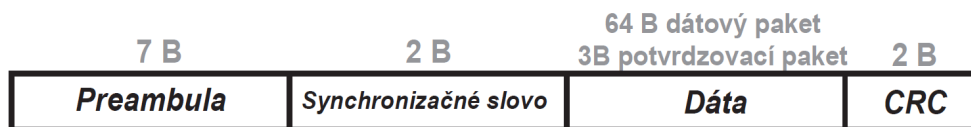


Obr.3 Cieľový hardvér

### C. Hlavné funkcie softvéru

Na bezdrôtový prenos dát pre radarovú sieť je kladená hlavne požiadavka zabezpečenia prenosu pre ochranu paketov pred stratou. Pri bezdrôtovej komunikácii dochádza vplyvom prostredia a rušení k strate dát. Program, ktorého úlohou je prenos zabezpečiť musí byť teda

navrhnutý tak, aby kontroloval prenos a v prípade potreby znovu poslal stratené pakety. Proces prenosu dát je softvérovým realizovaný podobne ako prenosový protokol TCP-IP. Jedná sa o zabezpečenie prenosu potvrdzovaním prijatia paketu alebo skupiny paketov. To sa realizuje odosielaním spätnej informácie o paketoch prijatých v prijímači. Vysielač následne realizuje prípadné opätovné odoslanie stratených paketov. Po úspešnom ukončení posielania sa proces prenosu opakuje. Výsledný paket posielaný rádiovými modulmi je zobrazený na obr. 4.



Obr.4 Štruktúra výsledného preneseného paketu

#### D. Kontrola spojenia

Systém komunikácie je navrhnutý tak, aby boli dáta prenášané, iba ak je vytvorený fungujúci bezdrôtový prenosový kanál. To znamená, že pri inicializácii komunikácie si obe strany potvrdia fungujúcu linku. Tá je využívaná po dobu prenosu a prenos je úspešne ukončený uzatvorením linky - spojenia. Toto riešenie bolo zvolené pre jednoduchšiu kontrolu prenosu. Prenos dát je totiž možný iba po plne fungujúcom prenosovom kanáli. Vytvorenie bezdrôtového spojenia je realizované skenovaním frekvenčného pásma na prítomnosť rušivých signálov, voľbou vhodných frekvenčných kanálov na komunikáciu a úspešnom spojení na dohodnutých kanáloch. Ukončenie spojenia je signalizované rozdielnou hlavičkou posledného dátového paketu, ktorá indikuje ukončenie spojenia. Spojenie vysielačieho programu v počítači po linke USB s vysielačom dát je uzavreté až po úspešnom spojení rádiových modulov.

#### E. Voľba vhodných frekvenčných kanálov

Pre zabezpečenie efektívnej komunikácie je potrebné voliť frekvenčné kanály, ktoré neobsahujú vysoký šum. To umožňuje dosiahnuť pri komunikácii vysoký pomer signál – šum. Pomer S/Š nie je v moduloch Silabs [5] priamo meraný, je totiž dosť problematické ho zmerať bez informácie čo je šum a čo užitočný signál. Pre voľbu frekvenčných kanálov sme teda hľadali adekvátnu náhradu. Tou je meranie celkovej sily prijatého signálu v moduloch Silabs označované ako parameter RSSI. Moduly merajú na vstupe silu prijatého signálu. Nakoľko sa meranie realizuje za kanálovým filtrom zodpovedajú získané informácie o sile signálu v kanále, na ktorý je prijímač naladený. Ak využijeme meranie v čase, kedy nedochádza k vysielačaniu na danom frekvenčnom kanáli je celkový prijatý signál iba šum. Toto meranie využijeme na určenie vhodných kanálov pre komunikáciu.

#### F. Nadviazanie rádiového spojenia

Jedným z problémov bezdrôtového prenosu je problematika vytvorenia vhodných podmienok pre nadviazanie spojenia. V našom systéme je nadviazanie kvalitného spojenia riešené skenovaním frekvenčného pásma a následne výberom vhodných frekvenčných kanálov pre prenos dát. Skenovanie frekvenčného pásma sa realizuje v prvom kroku postupným ladením jedného z rádiových prijímačov na každý frekvenčný kanál a následne opakovaným načítaním a sprimerovaním načítaných hodnôt parametra RSSI. Nameraný parameter sa následne uloží do poľa. V ďalšom kroku sa pole zoradí podľa veľkosti, pričom sa pracuje aj s druhým poľom, v ktorom sledujeme pozíciu nameraného RSSI parametra pre každý kanál. Po zoradení dostaneme pole s informáciami, ktoré kanály sú najvhodnejšie, majú najnižší RSSI, teda šum alebo rušenie na kanáli. Následne sa zvolia vhodné kanály s vhodným odstupom a informácia o vybraných kanáloch sa odošle na nastavených začiatkových kanáloch do koncových uzlov. Tie potvrdia prijatie kanálov a následne dôjde k preladeniu všetkých zariadení na dohodnuté kanály. Po úspešnom spojení na nových kanáloch je prenosová linka pripravená na prenos dát a koncový uzol potvrdí pripravenosť zdroju dát na linke USB.

#### ZÁVER

Úspešne sme navrhli a zrealizovali bezdrôtový prenosový systém pre prenos dát medzi jedným koncovým uzlom radaru a koncentrátorom dát. Navrhnutý systém je popísaný v bakalárskej práci [6]. Navrhnutý systém sme testovali pre overenie spoľahlivosti linky a možnosti jej ďalšieho rozvoja. Všetky komponenty systému máme pripravenými funkciami úplne pod kontrolou a je teda možné jednoducho a efektívne systém ďalej rozširovať alebo prípadne meniť cieľovú funkcionalitu. Dodržali sme podmienku, ktorú sme si zadali a to vytvorenie modulárneho systému, ktorý je možné jednoducho využiť pre iné bezdrôtové moduly, iný riadiaci mikroprocesor alebo pre pridanie ďalších rádiových modulov. V súčasnom

stave je systém vytvorený ako obojsmerná dátová linka medzi dvoma zariadeniami, ktorá je navrhnutá na prenos dát jedným smerom - z koncového uzla do koncentrátora.

Bezdrôtový systém sme testovali a výsledky naznačujú, že zvolené riešenie vyhovuje požiadavkám na rýchlosť prenosu a zabezpečenie bezstratovosti prenosu. Testy naznačili aj nedostatky navrhnutého riešenia a smer ďalšieho vývoja systému. V zabezpečení ochrany dát pred stratou vidíme ešte možnosti na zlepšenie zabezpečenia prenosu. Jedným z možných riešení zabránenia stratám pri prenose by bolo predĺženie preambuly, čo by ale malo za následok zníženie prenosovej rýchlosti.

Dôraz sme kládli na vytvorenie funkcií pre jednotlivé činnosti tak aby bolo možné tieto funkcie ďalej využiť a nemusí sa teda jednáť o výsledné vývojové štádium zariadenia ale je pravdepodobný nasledujúci rozvoj. V ďalšom rozvoji prehodnotíme výsledky testov a zakomponujeme úpravy.

Projekt je možné v budúcnosti ďalej rozvíjať a to hlavne zväčšením prenosovej kapacity linky zmenou modulácie modulov a pridaním ďalších modulov bezdrôtovej komunikácie pre vytvorenie komunikačného systému pre radarovú sieť so 4 radarmi. Ďalší rozvoj plánujeme práve týmto smerom, pričom s využitím existujúcich funkcií bude možné projekt jednoducho rozšíriť o potrebnú funkcionálnosť podľa ďalších požiadaviek.

#### POĎAKOVANIE

Tento článok bol vytvorený v rámci realizácie projektu UWB senzorové siete krátkého dosahu na detekciu, lokalizáciu a sledovanie pohybujúcich sa osôb (VEGA projekt č 1/0563/13) a realizácie projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] G. O. Young, J. Sachs 2003. M-Sequence Ultra-Wideband-Radar: State of Development and Applications. September 2003. Dostupné na internete: <http://www-emt.tu-ilmenau.de/EMTPub/uploads/pdf/3289561733U42I.pdf>
- [2] Electronic communications Committee 2014. ERC Recommendation 70-03. February 2014. Dostupné na internete: <http://www.ero-docdb.dk/docs/doc98/official/pdf/rec7003e.pdf>
- [3] ARM Ltd. Cortex-M4 Processor. 2014. Dostupné na internete: <http://www.arm.com/products/processors/cortex-m/cortex-m4-processor.php>
- [4] STMicroelectronics. STM32F4DISCOVERY. 2014. Dostupné na internete: <http://www.st.com/web/catalog/tools/FM116/SC959/SS1532/PF252419>
- [5] Štefan Hedvig, Viackanálový prenos dát vo frekvenčnom pásme SRD. Bakalárska práca, 2014.
- [6] Silicon Laboratories. SI4362. 2013. Dostupné na internete: <https://www.silabs.com/Support%20Documents/TechnicalDocs/SI4362.pdf>

# Detekcia, sledovanie a identifikácia pohyblivých cieľov vo videosekvenciách s využitím časticových filtrov

<sup>1</sup>Michal SZAKALA, <sup>2</sup>Ján VALISKA, <sup>3</sup>Stanislav MARCHEVSKÝ

<sup>1,2,3</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky Technická univerzita v Košiciach, Slovenská Republika

<sup>1</sup>michal.szakala@student.tuke.sk, <sup>2</sup>jan.valiska@tuke.sk, <sup>3</sup>stanislav.marchevsky@tuke.sk

**Abstrakt** – Cieľom článku je návrh algoritmu na zachytenie pohybujúcich sa objektov vo videosekvenciách a ich sledovanie pomocou časticového filtra. Detekcia a segmentácia jednotlivých objektov môže byť vykonávaná rôznymi metódami. Jednou z nich je metóda optického toku, ktorá bola použitá pri zachytení pohybujúceho sa objektu. Jednotlivé segmentované objekty sú uložené do databázy a môžu byť uskutočnené rôzne metódy spracovania obrazov pre uľahčenie rozpoznávania zvoleného objektu. Testovanie algoritmu bolo vykonané pomocou statickej kamery a pomocou kamery polohovanej dvojicou servopohonu. Algoritmus je schopný sledovať viaceré objekty, ktoré sú rozlíšiteľné ich trajektóriou pohybu. Trajektória umožňuje mať prehľad o predchádzajúcich polohách a prípadne aj predpovedať budúce pozície objektov.

**KLúčové slová** – Optický tok, časticový filter, sledovanie, identifikácia objektov

## I. ÚVOD

gentné, že dokážu zaznamenať pohyb, resp. sledovať danú osobu. Väčšina kamerových systémov sú statické, v súčasnosti začali sa používať kamery, ktoré majú širšie zorné pole a to vďaka možnému otáčaniu kamery. Takéto sledovanie bolo nasimulované pomocou navrhnutého algoritmu.

Navrhnutý algoritmus používa na detekciu metódu optického toku, ktorý je výhodný oproti metóde sledovania na základe farby v zmysle nezávislosti na zmene intenzity farby počas snímanej scény, ktorá znižuje presnosť sledovania. Na sledovanie je použité časticové filtrovanie, ktoré z dôvodu rýchlosti súčasných počítačov sa dostalo do popredia oproti Kalmanovým filtrom. Kalmanov filter aj časticový filter patria do skupiny bayesovských filtrov. Rozdiel medzi Kalmanovým filtrom a časticovým filtrom spočíva v tom, že pri použití Kalmanovho filtra funkcie hustoty pravdepodobnosti musia mať gaussovské rozloženie. V tomto prípade Kalmanov filter je optimálnym bayesovským filtrom. V mnohých prípadoch však rozdelenia pravdepodobnosti sú nelineárne a negaussovské. V tomto prípade prichádza do úvahy časticový filter. Algoritmus zároveň umožňuje ukladať sledované objekty do databázy za účelom identifikácie.

## II. OPTICKÝ TOK

Podstatou optického toku je zaznamenať pohyb v obraze. Samotný pohyb však v trojrozmernom reálnom svete je iný ako v dvojrozmernom obrazovom priestore. Vo väčšine prípadov pohybujúci sa objekt ukáže konštantnú intenzitu jasovej charakteristiky v rámci priebehu z jednej snímky na druhú.

Optický tok je rozloženie zdanlivých rýchlostí pohybu jasu v obraze. Nesúvislosti optického toku môžu pomôcť v segmentácii obrazu na regióny, ktoré korešpondujú rôznymi objektmi. Najčastejšie používané metódy určenia optického toku sú diferenčné metódy, prispôbenie na základe regiónov, metódy založené na energii a metódy založené na fáze. Metódy sú podrobnejšie popísané a porovnané v [1].

Napriek svojím rozdielom, metódy pozostávajú z troch krokov:

- 1) prefiltrovanie alebo vyhladenie s dolnopriepustným filtrom resp. pásmovým priepustom na extrakciu vybranej signálovej štruktúry a za účelom zvýšenia pomeru signál/šum.
- 2) extrakcia základných meraní, napr. derivácie na meranie rýchlosti alebo lokálne korelačné plochy
- 3) integrácia meraní na získanie dvojrozmerného poľa toku, ktorá často zahŕňa predpoklady plynulosti poľa toku

### III. ČASTICOVÉ FILTRE

Hlavná myšlienka časticového filtrovania je reprezentácia posteriórnej pdf funkcie skupinou diskretných vzoriek známych ako častice. Názov častice vyplýva z diskretnéj povahy reprezentácie pdf. Každá častica reprezentuje hypotézu stavu. Častice sú vykresľované náhodne z priórnej hustoty. Po vykresľovaní sa častice pohybujú na základe stavového modelu. Každá častica je charakterizovaná váhou, ktorú udáva merací model. Váha charakterizuje kvalitu danej častice. Pravdepodobné častice majú vyššie váhy, menej pravdepodobné nižšie.

Posteriórna funkcia hustoty rozdelenia pravdepodobnosti je konštruovaná rekurzívne skupinou náhodných vzoriek s príslušnými váhami  $\{x_t^{(i)}, w_t^{(i)}; 1, \dots, N\}$ , kde  $N$  je počet častíc.

Kroky rekurzívneho Bayesovského filtrovania, t.j. predikcia a aktualizácia sú riešené nasledovne:

- Predikcia: každá častica  $x_t^{(i)}$  sa vyvíja nezávisle podľa stavového modelu s prídavným šumom na simuláciu neznámych vyskytujúcich rušení. Aproximovaná priórna pdf funkcia je udávaná ako:

$$p(x_t) \approx \frac{1}{N} \sum_{i=1}^N \delta(x_t - x_t^{(i)}) \quad (1)$$

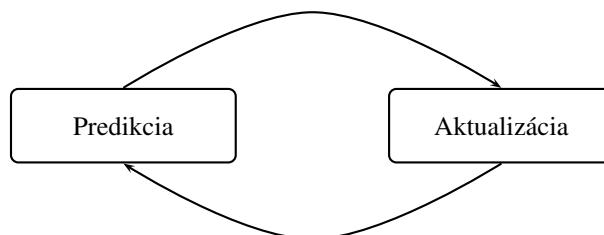
- Aktualizácia: váhy častíc sú nastavené na základe najnovšieho merania pomocou meracieho modelu. Posteriórna pdf funkcia je udávaná ako:

$$p(x_t | z_{1:t}) \approx \sum_{i=1}^N w_t^{(i)} \delta(x_t - x_t^{(i)}) \quad (2)$$

kde

$$w_t^{(i)} = \frac{\zeta(x_t | x_t^{(i)}) p(x_t^{(i)} | x_{t-1}^{(i)})}{q(x_t^{(i)} | x_{t-1}, z_t)} \quad (3)$$

Hustota  $q(x_t^{(i)} | x_{t-1}, z_t)$  je použitá v prípade, keď je nemožné získať vzorky z priórnej hustoty. Táto hustota má názov váhová (dôležitá) hustota. [2] Myšlienka dôležitého vzorkovania môže byť použitá na generovanie náhodných premenných zo zložitých hustôt. [3]



Obr. 1: Zovšeobecnený algoritmus časticového filtrovania

Časticový filter pri dôležitom vzorkovaní vykazuje problém degenerácie, kde po niekoľkých iteráciách všetky častice okrem jednej budú mať zanedbateľné váhy, t.j. ich rozptyl prudko klesá.

Na riešenie tohto problému existujú rôzne metódy, prakticky najdôležitejšie sú 2 hlavné metódy: [4]

- 1) vhodná voľba dôležitej hustoty
- 2) prevzorkovanie

Prevzorkovanie je metóda na zníženie efektu degenerácie pri dôležitom vzorkovaní. Hlavná myšlienka prevzorkovania je eliminovanie častíc, ktoré majú nízke váhy a sústrediť sa na častice s vyššími váhami.

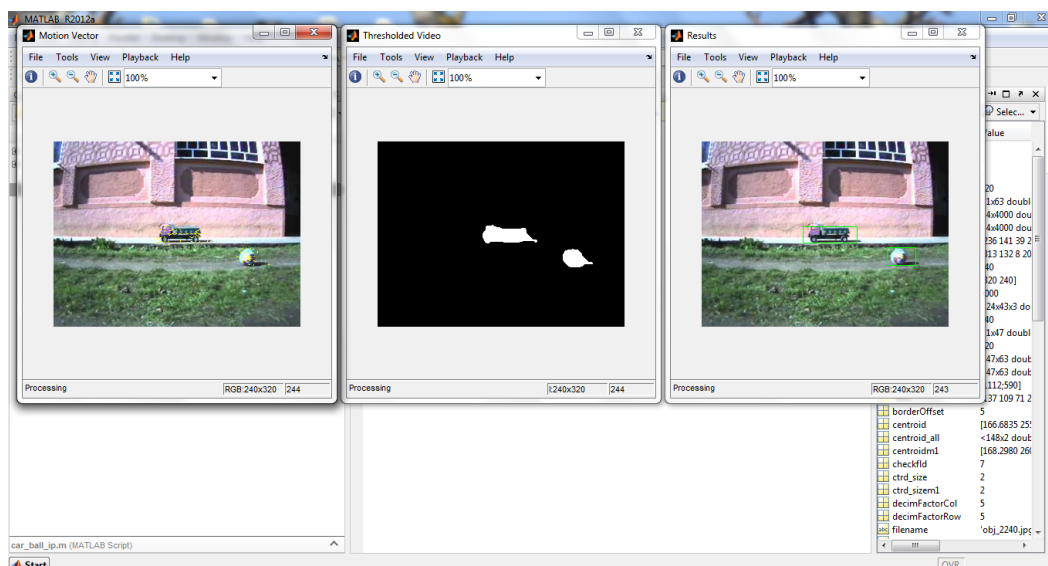


#### IV. SIMULÁCIE

Uvažované prípady sledovania na testovanie algoritmu boli:

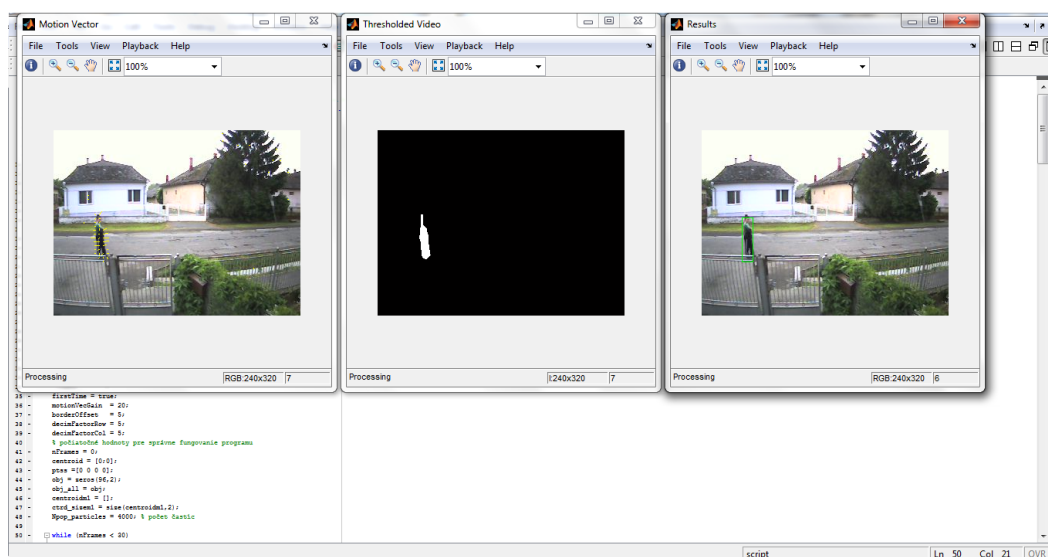
- Sledovanie vo vopred uloženom videu
- Sledovanie v reálnom čase
- Sledovanie s otáčaním kamery

Algoritmus sledovania vo vopred uloženom videu umožní sledovanie viacerých objektov súčasne. Pri sledovaní sa používa jeden filter na sledovanie všetkých objektov. Nízke počty objektov nemajú vplyv na plynulosť procesu sledovania aj keď je použitý len jeden filter. Podľa t'žísk objektov je vykresľovaný rámček na základe 500 vybratých častíc s najvyššími váhami po operácii prevzorkovania. Počas sledovania sa ukladajú snímky jednotlivých objektov do databázy, ktorá je vygenerovaná na základe počtu objektov. Po vytvorení databázy je možná implementácia identifikácie.



Obr. 2: Sledovanie dvoch objektov

Priebeh sledovania v reálnom čase ukáže Obr. 3, kde bol sledovaný jeden pohyblivý objekt pomocou 4000 častíc v rozlíšení  $320 \times 240$ . Jednotlivé okná predstavujú rovnaké funkcie ako pri sledovaní vo vopred uloženom videu.



Obr. 3: Sledovanie v reálnom čase

Sledovanie s otáčaním kamery je užitočné pri sledovaní osôb, kde je možné sledovať osobu dlhšiu dobu ako pri sledovaní so statickou kamerou. Simulácia bola vykonaná vo vonkajšom prostredí s pohybom jedného sledovaného objektu. Pri sledovaní sa používali 4000 častíc. Proces sledovania je ten istý ako pri sledovaní vo vopred uloženom videu s rozdielom nutnosti výskytu len jedného objektu pri sledovaní. Kamera sa otáča súhlasne s pohyblivým objektom.

## V. ZÁVER

Článok bol venovaný tematiky detekcie, sledovania a identifikácie pohybujúcich sa objektov. Z možných metód detekcie objektov bola použitá diferenčná metóda optického toku. Ako prostriedok na sledovanie bolo použité časticové filtrovanie, ktoré i napriek svojej výpočtovej náročnosti preukázal vysokú presnosť a pomerne rýchle sledovanie. Navrhnutý algoritmus je najúčinnjší pri sledovaní vo vopred uloženom videu, kde je možné sledovanie viacerých objektov súčasne. Simulácie v reálnom čase a s otáčaním kamery boli urobené z dôvodu vyskúšania algoritmu v reálnom prostredí prostredníctvom sieťovej kamery. Pri správnej implementácii algoritmus môže predstaviť základ k ďalšiemu rozvíjaniu sledovania resp. identifikácie. Možné rozvíjanie programu je implementácia s knižnicou OpenCV pre dosiahnutie vyšších rýchlostí pri spracovaní v reálnom čase. Ďalšia možnosť detekcie je využívanie hybridnej techniky napr. kombináciou optického toku a farebného histogramu.

## POĎAKOVANIE

This work was supported by the Slovak Cultural and Educational Grant Agency under contract No. 010TUKE- 4/2012 and by the Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic under the contract No. 1/0563/13 and the Slovak Research and Development Agency under the contract No. APVV-0404-12.

## LITERATÚRA

- [1] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vision*, vol. 12, no. 1, pp. 43–77, Feb. 1994. [Online]. Available: <http://dx.doi.org/10.1007/BF01420984>
- [2] K. K. Ng and E. J. Delp, "New models for real-time tracking using particle filtering," pp. 72 570B–72 570B–12, 2009. [Online]. Available: <http://dx.doi.org/10.1117/12.807311>
- [3] T. B. SCHÖN, "Solving nonlinear state estimation problems using particle filters – an engineering perspective," Division of Automatic Control, Linköping University, Sweden, Tech. Rep. LiTH-ISY-R-2953, May 2010.
- [4] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *Trans. Sig. Proc.*, vol. 50, no. 2, pp. 174–188, Feb. 2002. [Online]. Available: <http://dx.doi.org/10.1109/78.978374>

# Universal model of operating systems migration

*Liberios Vokorokos, Michal Ennert and Eva Chovancová*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

liberios.vokorokos@tuke.sk, michal.ennert@tuke.sk, eva.chovancova@tuke.sk

**Abstract** — In this paper, we describe design and creation of a universal model of operating systems migration. The created model should serve as an aid to standardization, planning and implementation of migrations. Paper deals with the problem of migration of operating systems and migration processes. Several existing solutions are analyzed and compared. Universal model is built on common characteristics of selected types of migrations. The model includes diagrams and procedures that are necessary to realize the implementation of migration. Testing is realized by implementing the selected migration on the basis of created model. After successful completion of the implementation mistakes and shortcomings are summarized and the designed universal model is modified.

**Keywords** — operating systems migration, operating systems deployment, universal model

## I. INTRODUCTION

Problem during change of the operating system may occur, if there is a domain, where end devices are already set to certain modes of operation. This means, that it is impossible to suspend the running machines and change their types of operating systems without affecting organization operability. The solution can be migration with minimal impact. Proper design, testing, implementation and support can reduce costs, increase stability, security and uniformity of whole system. Universal model described in this article can provide solutions leading to successful configuration and migration of target devices. To create solution design, it is necessary to describe functionality and principles of individual components, protocols and services that are essential for migration. The result of migration should be higher safety, homogeneity, easier manageability, reliability and lower costs. There are many types of migration, and each of them can have a different other variations. Taking a closer look, we can see a certain similarity between those migrations, such as on the level of the network architecture or on the software one. The idea is to summarize common components, processes and protocols to a more abstract level, based on which it is possible to construct a universal model of migration, on which will be applied common procedures for changing operating systems in organizations.

## II. ANALYSIS

The process of downloading, installing and updating software with the minimum participation of the user (the user just selects which operating system or application wants to install or update) is called software deployment. Software deployment is specifically defined in [1] and says, that it is a collection of related activities which form the life cycle of software deployment, which is performed both on the issuer and consumer side. Software deployment life cycle describes author in [2], and divides it into the following processes: release, installation, activation, updating, adaptation and uninstallation. As mentioned in [3], a model of software deployment may specify policies of process control and type of scheduling and synchronization. It is appropriate that the control of the deployment process is automatic or semi-automatic at most, whereas the involvement of the user in the process should be as small as possible. Author subdivided planning and synchronization the incremental, continuous and "Big-Bang" type. With the software deployment is related discipline of configuration management, which is in [4]

is defined as the discipline that manages the evolution of the software product during its development phase. As an example of the software covered by this discipline we can mention System Center Configuration Manager (SCCM) from Microsoft or Novell Zenworks. We can divide software deployment into two areas, namely application deployment and operating system deployment. Under the application deployment this paper understands the distribution and installation of programs, their updates and patches. Many programs allow silent installation, which in practice means, that the selected application will be installed without or with minimal user interaction. This type of installation can be configured using the silent switches. Deployment of applications on the remote workstations in combination with their silent installation author describes as one of the approaches in [5]. In order to successful migration, it is necessary to solve several problems, such as lack of receiver, dynamic shutdown/startup, device state abstraction and so on. These problems and their solutions describes author in [6].

### III. RELATED WORK

We describe several existing solutions for the design of a universal model of migration of operating systems. From these solutions are then selected common or similar infrastructure components, procedures, protocols, software solutions etc. Based on these common elements the universal model of migration of operating systems is created. For mentioned types of migration the presence of a working domain and DNS server is essential.

#### A. Migration from Windows XP to Windows 7

Selected type of migration is based on the project in Volkswagen Slovakia, Inc. Our solution considered including auto-deployment software to increase process automation. Preparation of migration consisted of several steps. This includes preparing deployment server for the Microsoft Deployment Toolkit 2010, setting up Windows Deployment services, creating an installation image from a reference computer, creating an answer file, distribution share, deployment point, setting up the DHCP server etc. Basic architecture required to migrate from Windows XP to Windows 7 is shown on Fig. 1.

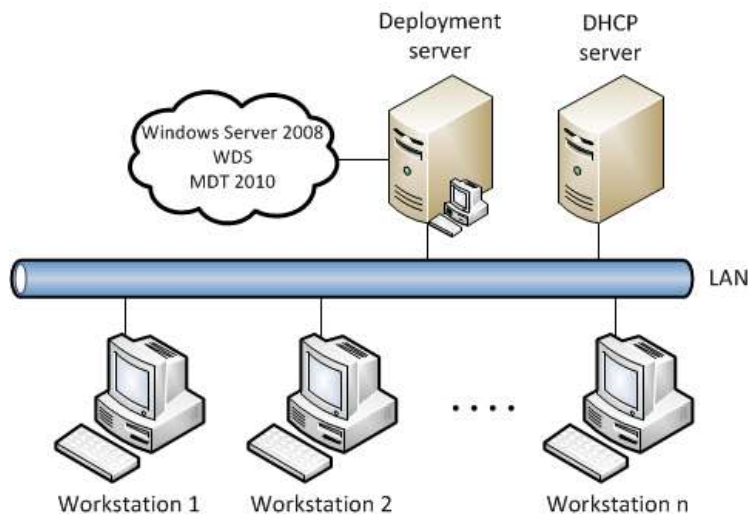


Fig. 1 Basic architecture required to migrate from Windows XP to Windows 7

#### B. Migration from Windows XP/7 to Ubuntu Desktop Edition

Another type of migration chosen for the creation of an universal model is based on the solution used in the Skalica City Hall. Similar to the previous case, for proper functionality of the deployment installation and configuration of specific components is required. For this case was chosen a tool, that has an integrated elements such as TFTP, configuration management etc. Selected software is Cobbler, which also serves to automatic deployment and updating of software. Similar to the previous migration, even in this case, answer file is necessary for automatic installation. Cobbler has an integrated tool Cheetah, which uses Kickstart templating. Configuration of DHCP server, which provides information needed to PXE boot to the workstations, is also necessary. Fig. 2 shows the basic architecture needed for this type of migration. The basic concept of architecture is the same as in the previous type of migration.

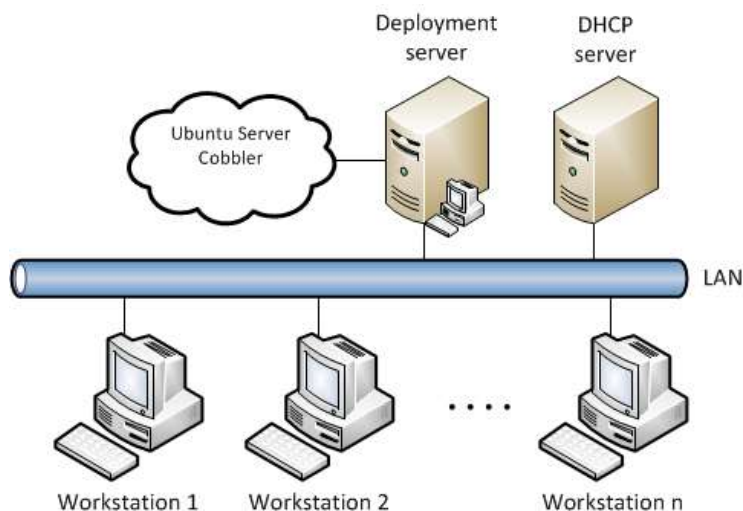


Fig. 2 Basic architecture required to migrate from Windows XP/7 to Ubuntu Desktop Edition

### C. Other types of migrations

Example can be switching from Linux distribution to Windows, which is sometimes referred as a reverse migration. Another, more specific, is optional migration with a choice of final operating system. The idea is that the user can select the operating system to be installed on his workstation. We can use the previously mentioned principles of migrations. However, the problem may be a lack of support images with different operating systems core in the Microsoft Deployment Toolkit and Windows Deployment Services. One approach to eliminate this problem is the presence of redundant deployment server. We can see migration with redundant deployment server on Fig. 3.

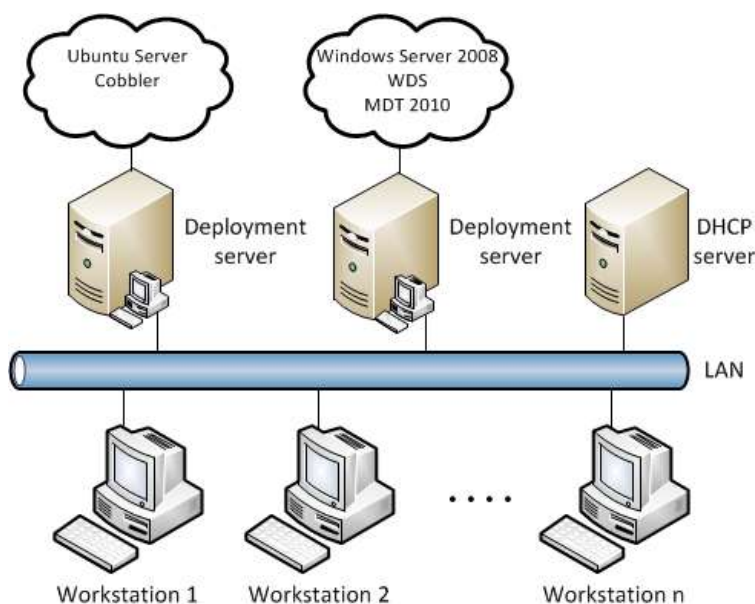


Fig. 3 Basic architecture needed for the optional migration

Last mentioned type of migration is between two Linux platforms. Since all Linux distributions using the same Linux kernel, this type of migration is technically the simplest of solutions mentioned.

## IV. UNIVERSAL MODEL OF MIGRATION

Based on these types of migration architectural elements, settings, and procedures can be generalized. Assuming that there is already a domain and configured DNS server, DHCP server and the domain management server, we can design a deployment diagram that describes

individual basic components required for software. General deployment diagram is shown in Fig. 4.

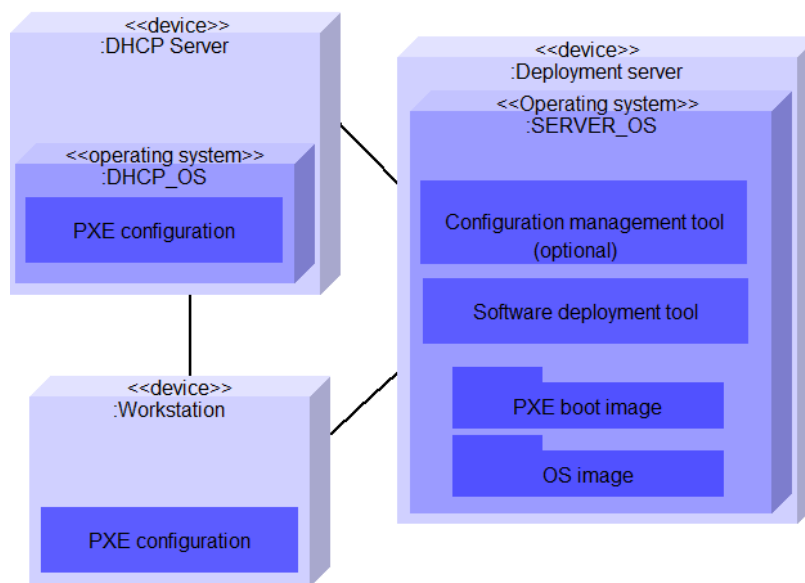


Fig. 4 General deployment diagram

Next, model describes the steps necessary to the migration implementation:

- 1. Server OS migration** is possible automatically according to [7] and [8], but this process can be quite complex, so it is more appropriate to deploy of a redundant server. After operating system installation, it is necessary to configure the server to provide services as well as the original server. If this condition is met, we can disconnect the original server and redundant server takes over its role. This step is optional.
- 2. Architecture preparation** involves the preparation of infrastructure for software deploying. First, it is necessary to configure deployment server. Then it is possible to install and configure tool for software deployment. If necessary, installation and setup of configuration management tool is possible. For automatic installation of the operating system it is necessary to create the answer file. If there is no image of the operating system, it needs to be prepared and created. If all previous steps are met, we need to integrate these images into the software deployment tool. DHCP server configuration is also required.
- 3. Data and user settings backup.** Some tools offer along with the operating system migration also data and settings migration. For example, Microsoft offers a User State Migration Tool, which can be combined with Windows Deployment Services. However, the solution can also be a backup to network drives, removable media, another disk partition and so on.
- 4. Workstation configuration.** Network boot must be enabled in BIOS settings. It is also advisable to check the SATA operation mode, as it should be consistent with the requirement of settings of the operating system to be installed on the station. Sometimes it is necessary to enable NIC boot with PXE.
- 5. Workstation migration.** The last step is the migration of workstations. After we install the operating system itself, depending on organizational requirements and settings, it may be required to install additional software, update some specific settings etc. After creating a universal model, implementation of migration for the testing purpose was realized. We can see architecture of implementation on Fig. 5.



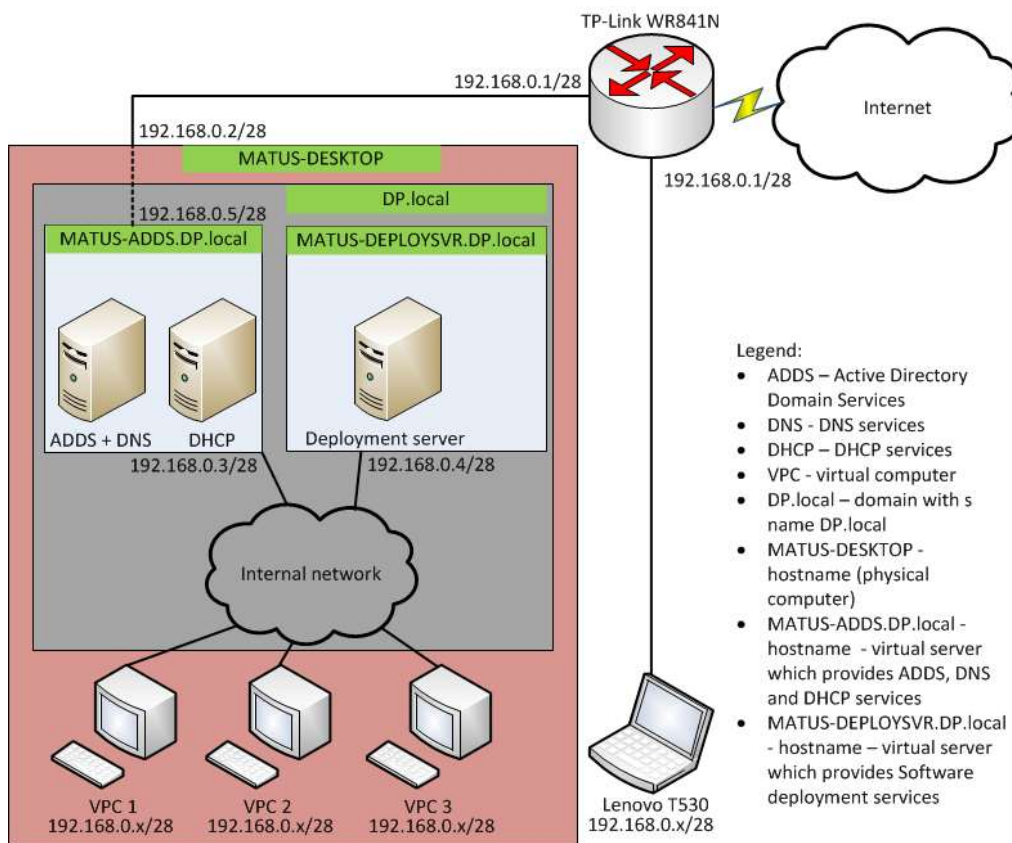


Fig. 5 Architecture of implementation infrastructure

We used Windows Server 2008 R2 Enterprise as OS for deployment server. The implementation used the Microsoft Deployment Toolkit 2010, System Center Configuration Manager 2012, ImageX and few other tools. Data backup was necessary only from the physical laptop, since virtual machines didn't have an operating system. These data have been backed up on external media.

## V. TESTS AND RESULTS

Realized implementation of migration based on designed universal model confirmed applicability of described principles and procedures. No major errors which could lead to impossibility of realization of any step of universal model were detected during implementation and testing phase. A problem has occurred during migration OS of physical laptop, which didn't have SATA bus controller set to AHCI operation mode in BIOS by default. This issue would result in an installation failure. Settings have been changed manually, so this obstacle was eliminated. Changing this setting is also possible automatically by configuring the script and adding it to tasks sequence into the System Center Configuration Manager. Presence of network and disc drivers in the boot and installation images is necessary condition. If these drivers are not integrated with images, system can't communicate with the HDDs or servers, which will lead to installation failure again. Given implementation did not deal with the migration of operating systems with 64-bit architecture. But as a separate image of this architecture would have to be created for each type of workstation once again, it is unlikely that this migration could be not successful. The solution was implemented in a predominantly virtual environment. Implemented was only migration to Windows 7. Existence of a configured DHCP server, DNS server, domain management server and also the existence of the domain itself are essentially conditions for this model. For better illustration it is appropriate to enlarge designed deployment diagram. This modified diagram is shown in Fig. 6.

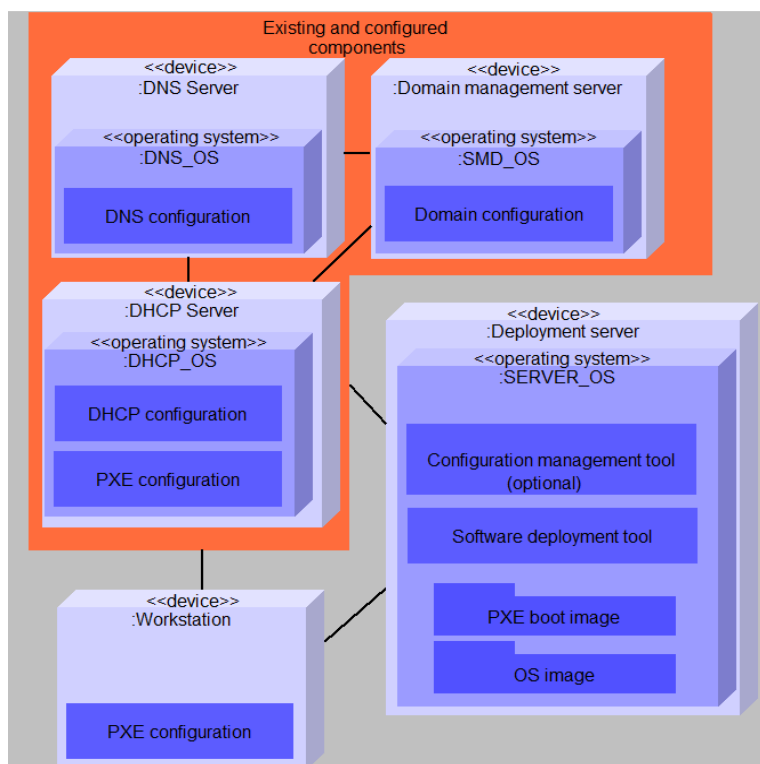


Fig. 6 Modified general deployment diagram of implementation infrastructure

## VI. CONCLUSION

This paper describes the basic concepts, definitions and procedures that are necessary to better understand the procedural steps associated with migration of operating systems. It also introduces and describes created universal model of operating systems migration. Using this model it is possible to standardize and simplify the tasks associated with planning, implementation and testing a migration of operating systems. Since the implementation of the migration process is a lengthy process that in the real environment may take several months, the implementation of other types of migration based on the established model was not realized and may be the subject of further research. The created model can serve as a useful tool for administrators and users involved in the process of migration of operating systems.

## ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and KEGA 008TUKE-4/2013 Microlearning environment for education of information security specialists. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] HALL, R. S., HEIMBIGNER, D., WOLF, A. L.: A cooperative approach to support software deployment using the software dock. Proceedings of the 21st international conference on Software engineering. ACM. May 1999. p. 174-183.
- [2] DEARLE, A.: Software Deployment, Past, Present and Future. FOSE'07: 2007 Future of Software Engineering. IEEE Computer Society. Washington, DC. 2007. ISBN 0769528295. p. 269-284.
- [3] COUPAYE, T., ESTUBLIER, J.: Foundations of Enterprise Software Deployment. Software Maintenance and Reengineering. IEEE Computer Society. Washington, DC. February 2000. ISBN 0769505465. p. 65-73.
- [4] VAN DER HOEK, A. Integrating configuration management and software deployment. Proceedings of the Working Conference on Complex and Dynamic Systems Architecture (CDSA 2001). December 2001.
- [5] GOMBERG, M., EVARD, R., STACEY, C.: A comparison of large-scale software installation methods on nt and unix. Proceedings of the Large Installation System Administration of Windows NT Conference (SAGE/USENIX). August 1998. p. 37.
- [6] KOZUCH, M. A., KAMINSKY, M., RYAN, M. P.: Migration without Virtualization. HotOS. 2009.
- [7] FEINLEIB, D. A., MORAN, B. K.: System and method for automated migration from Windows to Linux. U.S. Patent No 7,805,721. September 2010.
- [8] FEINLEIB, D. A., MORAN, B. K.: System and method for automated migration from Linux to Windows. U.S. Patent No 7,284,043. 2007.
- [9] VOKOROKOS, L., BALÁŽ, A., MADOŠ, B., RADUŠOVSKÝ, J.: Cloud file system. IEEE 17th International Conference on Intelligent Engineering Systems : proceedings : June 19-21, 2013, Costa Rica. - Budapest: IEEE, 2013 P. 323-326. - ISBN 978-1-4799-0828-8.



# Virtualizovaná laboratórna architektúra s využitím platformy od spoločnosti Microsoft

<sup>1</sup>Michal ONDIČ, <sup>2</sup>Roman MIHAL, <sup>3</sup>Iveta ZOLOTOVÁ

<sup>1,2,3</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach

<sup>1</sup>michal.ondic@student.tuke.sk, <sup>2</sup>roman.mihal@tuke.sk, <sup>3</sup>iveta.zolotova@tuke.sk,

**Abstrakt** — Článok je zameraný predovšetkým na popísanie jednotlivých virtualizačných a cloudových trendov v závislosti od súčasných technológií. Jednotlivé technológie sú rozanalyzované a vysvetlené vzhľadom k štyrom významným spoločnostiam, ktoré tieto technológie zaviedli a poskytujú ich pre potreby vytvorenia virtualizovaného pracoviska. Účelom riešenia v rámci laboratória bolo zrealizovať virtualizovanú architektúru, ktorá je založená na produktoch od spoločnosti Microsoft. Táto architektúra obsahuje prvky informačných a riadiacich systémov, pričom sa prihlíada na dosiahnutie vysokej dostupnosti a bezpečnosti.

**Príučové slová** — Virtualizácia, virtualizovaná architektúra, vysoká dostupnosť, cloudové technológie, hypervízor.

## I. ÚVOD

V rámci zabezpečenia sféry pohybujúcej sa v oblasti riadenia systémov, väčšina spoločností preferuje prístup, ktorý je založený na princípoch virtualizácie, z dôvodu šetrenia času, nákladov, údržby, ale aj ďalších vplyvov, ktoré majú určitú výhodu oproti spôsobom založených bez využitia virtualizačnej správy daného komplexu. Funkčnosť veľkých dátových centier je v súčasnosti založená na x64 serveroch v závislosti od zvolenej platformy daného riešenia.

Celá myšlienka je založená na prevádzke serverovej virtualizácie v rámci fyzických serverov, na ktorých bežia virtuálne stroje. Za pomoci virtualizácie je umožnené vykonávanie zložitej konfigurácie, testovania a riadenia viacerých aplikácií, ktoré sú určitým zoskupením v rámci vybranej platformy. V súčasnosti existujú určité spoločnosti, ktoré poskytujú jednotlivé platformy svojim zákazníkom k dispozícii pre zavedenie vlastnej virtualizovanej architektúry.

Najvýznamnejšími spoločnosťami, ktoré sa aktuálne zaoberajú a venujú virtualizačným a cloudovým trendom sú najmä VMware, Microsoft, Citrix a Oracle. [1]

## II. PREHLAD VIRTUALIZAČNÝCH A CLOUDOVÝCH TECHNOLOGIÍ

### A. Virtualizácia

V rámci tohto pojmu rozoznávame viacero druhov virtualizácie v závislosti od prístupu k daným prostriedkom. Môžeme rozlišovať virtualizáciu operačných systémov, aplikácií, riadenia, siete, hardvérovú, alebo softvérovú.

V intenciách tohto článku nás bude zaujímať najmä hardvérová virtualizácia (typy: plná, čiastočná, paravirtualizácia), ktorej cieľom je využiť jednotlivé zdroje hardvéru pomocou tenkej virtualizačnej vrstvy – hypervízora pre dosiahnutie určitého výkonu v rámci samostatných virtuálnych strojov bežiacich na tomto hardvéri.).

### B. Výhody virtualizácie

Zavedením virtualizácie v rámci určitej infraštruktúry pre zavedený objekt dostávame určité výhody plynúce na základe zvolenej technológie: [2]

- Zníženie prevádzkových nákladov na údržbu a školenie
- Zjednodušenie aplikačného a počítačového riadenia
- Zníženie licenčných nákladov

- Zvýšenie využitia hardvéru s cieľom zabezpečenia efektívnejšieho využitia zdrojov prostredníctvom konsolidácie serverov
- Posilnenie obchodnej kontinuity
- Zjednodušenie zálohovania a obnovy systému
- Rýchla návratnosť investícií
- Odstránenie problémov s nekompatibilitou aplikácií
- Zrýchlenie zabezpečenia servera
- Zníženie priestorových požiadaviek a spotrebovanej energie

### C. Hlavné požiadavky zavedenia virtualizovanej architektúry

Pri zavedení virtualizovanej architektúry je potrebné brať do úvahy požiadavky a doporučenia, ktoré sú základom kvalitnej, stabilnej, dostupnej a hlavne bezpečnej prevádzky virtualizačných technológií.

Pri budovaní infraštruktúry pre laboratórium, alebo obchodnú prevádzku je vhodné si uvedomiť nasledujúce skutočnosti: [3]

- Vybrať vhodnú technológiu – táto technológia sa vyberá pomocou požiadavky pre maximálne využitie virtuálnych počítačov, ktoré chceme mať k dispozícii. Pre tento krok je charakteristické zabezpečenie určitej úrovne služieb, ktorú môžeme dosiahnuť v závislosti od zdrojov daného servera.
- Určenie maximálnej novej zátáže fyzického servera – pri tejto zátäži musíme počítať s maximálnym efektívnym využitím výkonu procesora.
- Určenie minimálnej kapacity servera – berieme do úvahy najmä hardvérové vybavenie pri zohľadnení počtu a kapacity procesorov, diskov a operačnej pamäte RAM.
- Zohľadnenie veľkosti operačnej pamäte RAM – v závislosti od veľkosti organizácie a počtu virtuálnych strojov, ktoré budú vytvárané a využívané pre dané virtualizačné dátové centrum. V súčasnej dobe majú veľké spoločnosti k dispozícii fyzické servery, ktoré majú k dispozícii veľkosť pamäte RAM už od 320 GB a viac.

### D. Cloudové technológie

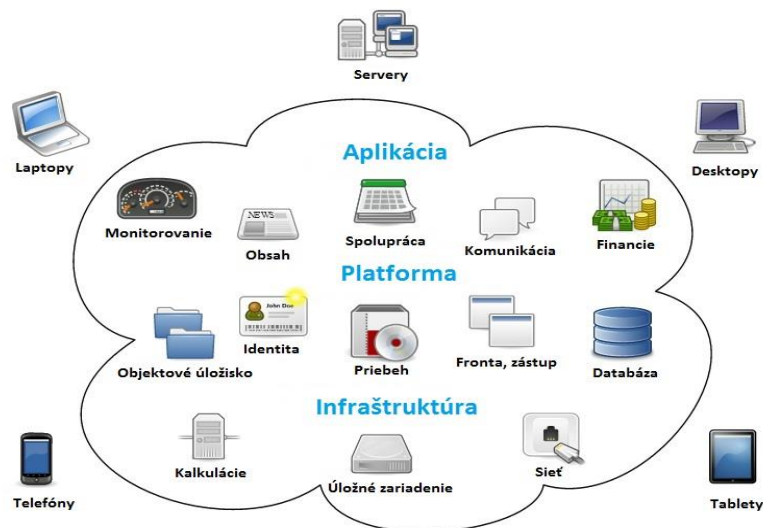
V rámci cloud computingu môžeme tieto technológie vnímať ako určitú formu, ktorá je neustále rozvíjaná v oblasti informačných a komunikačných technológií. O cloud computingu môžeme hovoriť ako o poskytovanej službe v oblasti IT, ktorá umožňuje zákazníkom pristupovať k aplikáciám, platformám a infraštruktúre rôzneho druhu v danom čase na diaľku. Hlavní poskytovatelia cloud computingových služieb majú dnes stovky až tisícky serverov, ktoré sa nachádzajú v masívnych výpočtových dátových centrách.

Cloudové služby sa vyskytujú v rôznych tvaroch a konfiguráciách:

- IaaS (Infraštruktúra ako služba)
- PaaS (Platforma ako služba)
- SaaS (Softvér ako služba)

V závislosti od spôsobu nasadenia danej služby užívateľom taktiež rozoznávame verejné, súkromné, komunitné a hybridné cloudy. Verejné cloudy poskytujú zoskupené služby poskytované pomocou siete, ktoré sú určené pre verejnosť, pričom pri súkromných ide o poskytovanie služieb pre konkrétnu spoločnosť, podnik, alebo organizáciu.

Komunitné cloudy sú tvorené službami, ktoré sú poskytované a zdieľané obmedzeným okruhom zákazníkov a sú riadené interne, alebo treťou stranou. Pri hybridných cloudoch ide o mix verejného a súkromného cloudu prepojeného určitými technológiami. [4]



Obr. 1 Cloud Computing

### III. SPOLOČNOSTI ZAOBERAJÚCE SA VIRTUALIZAČNÝMI TECHNOLOGIAMI

#### A. Spoločnosť ORACLE

Zahŕňa riešenie pre virtualizáciu serverov pomocou technológie (platformy) Oracle VM Server, ktorá je založená na x86 serveroch a spravovaná VM Managerom. K najvýznamnejším produktom patrí najmä Oracle databáza, Oracle Fusion Middleware, Oracle aplikácie a aplikačné klastre.

Pomocou nástroja Oracle VM môžeme ľahko a efektívne spravovať virtuálne stroje, kde pre nich môžeme zahrnúť najmä vytváranie, upravovanie, zdieľanie, správu a nasadenie v rámci hostovských (fyzických) serverov. Medzi podporované systémy tejto platformy patria Oracle Solaris, Linux, Red Hat Enterprise Linux a Microsoft Windows Server. [5]

#### B. Spoločnosť CITRIX

Je jednou z hlavných poskytovateľov virtualizačnej platformy, ktorá sa zaoberá vytváraním privátneho a verejného cloudu. Platforma s názvom XenServer je používaná najmä pre serverovú a desktop virtualizáciu. Jej hlavnou úlohou je virtualizovať pracovné úlohy a automatizovať procesy riadenia za zvýšenej flexibility a pružnosti v IT pri nízkych nákladoch.

Okrajovo je XenServer známy aj tým, že umožňuje prístup do Citrix vedeckého centra, odkiaľ môžeme získať informácie o nástrojoch pre konfiguráciu, optimalizáciu a riešenia nastaných problémov. [6]

#### C. Spoločnosť VMWARE

Táto spoločnosť si stále drží svoje prvenstvo v poskytovaní virtualizačných a cloudových technológií. Funguje na princípe virtualizačnej platformy VMware vSphere, pričom je využívaná viac ako 500 tisíc zákazníkmi po celom svete. Keďže je určená najmä pre fyzické servery zahŕňa v sebe hypervízor (tenká vrstva medzi hardvérom a softvérom) ESXi, ktorý bol vytvorený pre účely virtualizácie.

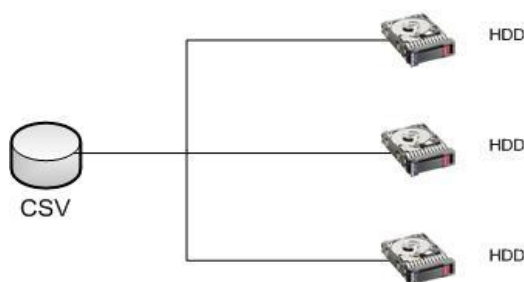
V súčasnosti existuje táto platforma vo verzii vSphere 5.5, pričom koncom roka bude už vydaná verzia vSphere 6.0. Celá správa pre fyzické servery s danou platformou je realizovaná pomocou nástroja vCenter Server.

Medzi technológie a funkcie súvisiace so zabezpečením dostupnosti a bezpečnosti patria najmä : High Availability (tvorba klastra vysokej dostupnosti), Fault Tolerance (predstavuje určitú odolnosť voči zlyhaniu virtuálneho stroja), Data Protection (zabezpečuje virtuálne prostredie pred stratou dát), vMotion (umožňuje migrovať virtuálne stroje), Replication (vytváranie kópií virtuálnych strojov na viacerých fyzických serveroch). [7] [8]

#### D. Spoločnosť MICROSOFT

Prvou virtualizačnou platformou od tejto spoločnosti, na ktorej sme mohli vykonávať virtualizáciu, bol Windows Server 2008 R2, ktorá už mala zabudovanú hypervízorovú virtualizačnú zložku Hyper-V. V súčasnosti je vhodnou platformou pre zavedenie prevádzky pre výpočtových centier platforma Windows Server 2012 v edícii Datacenter, ktorá umožňuje zaviesť až 64 fyzických serverov a 4000 virtuálnych strojov v rámci klastra vysokej dostupnosti.

Pre správu fyzických serverov a virtuálnych strojov bez zavedenia klastra slúži nástroj Hyper-V Manager, ktorý obsahuje základné nastavenia a funkcionality ako migráciu, replikáciu a pod. V rámci nástroja Failover Cluster Manager už ide o zabezpečenie vysokej dostupnosti pre virtuálne stroje, súborové servery, fyzické servery pomocou virtuálnych hard diskov, ktoré sú uložené na Cluster Shared Volume (CSV) diskoch. Tieto disky musia byť formátovo ako SAS disky, pričom pre zavedenie zdieľaného disku pre klaster sú potrebné minimálne 3 fyzické disky o kapacite minimálne 4 GB.



Obr. 2 Cluster Shared Volume

Komplexná správa je zabezpečovaná nástrojom System Center, ktorý vykonáva dohľad nad serverovými a klientskými časťami virtualizačného prostredia. V rámci neho je vykonávaná konfigurácia a automatizované riešenie problémov, ktoré nastali.

K najznámejším verziám System Center patria: [9] [10]

- App Controller, Data Protection Manager, Operations Manager, Service Manager, Unified Installer, Configuration Manager, Endpoint Protection, Orchestrator, Virtual Machine Manager

**Vysoká dostupnosť** môže byť v tomto nasadení charakterizovaná ako dostupnosť aplikácií, ktoré pracujú na virtuálnych strojoch, ak neberieme do úvahy na ktorom operačnom systéme pracujú. Hlavným jej účelom na tejto platforme je poskytnúť ochranu voči zlyhaniu hardvéru za pomoci nástroja Failover Cluster Manager. V rámci tohto nástroja je hlavnou úlohou zabezpečiť virtuálne stroje pred stratou dát s ktorými pracujú. [11]

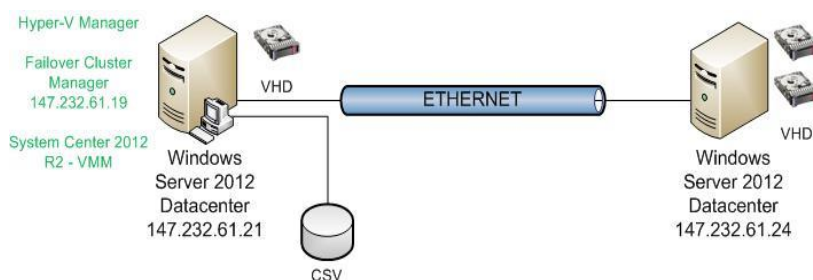
**Migrácia** je ďalšou z funkcií, ktorá môže byť vykonávaná v rámci Failover Cluster Managera. Rozoznávame dva druhy migrácie – Live migration (za behu virtuálneho stroja) a Quick migration (po vypnutí virtuálneho stroja). [12]

**NIC Teaming** bol taktiež využívaný pre realizáciu virtualizovanej architektúry. Pri tejto technológii ide o párovanie (spájanie) virtuálnych sieťových adaptérov daných virtuálnych strojov do tzv. uceleného konvergovaného switchu. V rámci platformy Windows Server 2012 Datacenter je umožnené spárovať až 32 virtuálnych sieťových kariet. [13]

#### IV. VIRTUALIZOVANÁ ARCHITEKTÚRA OD SPOLOČNOSTI MICROSOFT

Na obrázku 3 je realizovaná výsledná architektúra laboratória L509, ktorá bola realizovaná úspešne so všetkými hore uvedenými nástrojmi pre splnenie účelu vysokej dostupnosti na ktorý bola zavedená. Nástroje, ktoré boli inštalované na fyzickom serveri sú uvedené na obrázku vľavo zeleným písmom. Server vľavo s IP 147.232.61.21 bol hl. serverom pre vykonávanie správy, riadenia a úkonmi súvisiacimi so zavedením vysokej dostupnosti virtuálnych strojov.

Obidva fyzické servery komunikovali prostredníctvom domény, ktorá bola taktiež zavedená. Pre tieto servery bolo použité zdieľané úložisko vo forme Cluster Shared Volume diskov, pozostávajúce z troch fyzických diskov.



Obr. 3 Výsledná architektúra laboratória

Táto architektúra s využitím všetkých spomínaných virtualizačných technológií bola navrhnutá pre lepšie a efektívnejšie využitie fyzických serverov z pohľadu výpočtových virtualizovaných centier pri čo najnižších nákladoch na prevádzku.

## V. ZÁVER

V článku boli spomenuté niektoré základné pojmy dôležité pri virtualizácii a cloud computingu. Hlavnou úlohou bolo predstaviť významné spoločnosti zaoberajúce sa virtualizačnými technológiami a realizáciu návrhu virtualizovanej architektúry pozostávajúcej z nástrojov a funkcií konkrétne od spoločnosti Microsoft.

Dostupnosť bola dosiahnutá pomocou klastra vysokej dostupnosti, ktorý ju zabezpečil a centrálna správa bola riadená nástrojom System Center – Virtual Machine Manager. Zavedením výslednej architektúry do prevádzky boli splnené všetky kroky a požiadavky potrebné pre jej správne a efektívne fungovanie.

## POĎAKOVANIE

Táto publikácia je podporená z grantu projektov KEGA No. 021TUKE-4/2012 (50%) a VEGA - 1/0286/11 (50%).

## POUŽITÁ LITERATÚRA

- [1] *Server virtualization in manufacturing*, STRATUS Technologies. Dostupné na internete: <http://www.stratus.com/~media/Stratus/Files/Library/WhitePapers/ServerVirtualizationMfg.pdf>.
- [2] M. Tulloch with Microsoft Virtualization Teams, *Understanding Microsoft Virtualization Solutions: From the Desktop to the Datacenter*. Washington: Microsoft Press, 2010. Dostupné na internete: <http://download.microsoft.com/download/5/B/4/5B46A838-67BB-4F7C-92CB-EABCA285DFDD/693821ebook.pdf>.
- [3] D. Ruest, N. Ruest, *Virtualizace. Podrobný průvodce*. Brno: Computer Press a.s., 2010, s. 395, ISBN 978-90-251-2676-9.
- [4] *Information Economy report 2013. The Cloud Economy and Developing Countries*. s. 120, ISBN 978-92-1-112869-7. Dostupné na internete: [http://unctad.org/en/PublicationsLibrary/ier2013\\_en.pdf](http://unctad.org/en/PublicationsLibrary/ier2013_en.pdf).
- [5] *Oracle VM Server for x86 FAQ*. 2012. s. 5. Dostupné na internete: <http://www.oracle.com/us/026952.pdf>.
- [6] *Citrix XenServer: Industry-leading open source platform for cost-effective cloud, server and desktop virtualization*. 2013. s. 7. Dostupné na internete: [http://www.citrix.com/content/dam/citrix/en\\_us/documents/products-solutions/citrix-xenserver-industry-leading-open-source-platform-for-cost-effective-cloud-server-and-desktop-virtualization.pdf](http://www.citrix.com/content/dam/citrix/en_us/documents/products-solutions/citrix-xenserver-industry-leading-open-source-platform-for-cost-effective-cloud-server-and-desktop-virtualization.pdf).
- [7] M. Šika, *333 tipu a triku pro VMware*. Brno: Computer Press a.s., 2012, s. 276. ISBN 978-80-251-3659-1.
- [8] *Why Choose VMware for Server Virtualization: A Comparative Analysis for New Virtualization Customer*. 2013, s. 13. Dostupné na internete: [http://www.vmware.com/files/pdf/vmware\\_advantage.pdf](http://www.vmware.com/files/pdf/vmware_advantage.pdf).
- [9] J. Kelbley, M. Sterling, *Windows Server 2008 R2 Hyper-V: Podrobný průvodce administrátora*. Brno: Computer Press a.s., 2011, s. 387. ISBN 978-80-251-3286-9.
- [10] *Microsoft Hyper-V*. Dostupné na internete: <http://blogs.technet.com/b/technetczsk/p/microsoft-hyper-v.aspx>.
- [11] T. Olzak, *Virtualization: Hyper-V and High Availability*. TechNet Magazine, 2011. Dostupné na internete: <http://technet.microsoft.com/en-us/magazine/hh127064.aspx>.
- [12] J. Savill, *Windows Server 2012 Storage Live Migration*. 2013. Dostupné na internete: <http://windowsitpro.com/windows-server-2012/windows-server-2012-storage-live-migration>.
- [13] N. Sharma, *Working with NIC Teaming in Windows Server 2012*. 2014. Dostupné na internete: <http://www.serverwatch.com/server-tutorials/working-with-nic-teaming-in-windows-server-2012.html>.

# Použitie evolučného algoritmu pre optimalizáciu energie

<sup>1</sup>Viktor ŠOLTÉS, <sup>2</sup>Martin PALA

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>soltvikt@gmail.com, <sup>2</sup>martin.pala@tuke.sk

**Abstrakt** — Článok sa venuje analýze hydroponických systémov, genetických algoritmov a ich vzájomnému spojeniu pre potrebu optimalizácie hydroponického systému. V tomto článku sú popísané spôsoby pestovania hydroponickými systémami a načrtnutý úvod do automatizácie týchto systémov. Rozvrhovanie činnosti hydroponických systémov ako aj ich optimalizácia predstavuje multikriteriálnu úlohu, ktorej riešenie je veľmi náročné z časového hľadiska, preto tento článok popisuje prístup k riešeniu tohto problému. Ako základný prostriedok na optimalizáciu hydroponického systému bol použitý evolučný algoritmus. Výsledný program je navrhnutý pre optimalizáciu hydroponického systému používaného na Technickej univerzite v Košiciach. Navrhnutý systém dokáže úspešne optimalizovať hydroponický systém podľa parametrov, ktoré zadá užívateľ v podobe vstupných parametrov do programu. Program teda v konečnom dôsledku je možné použiť nie len na rozvrhovanie činnosti daného systému, ale zároveň umožňuje aj optimalizovať systém podľa konkrétnych požiadaviek.

**KLúčové slová** — pestovanie, hydroponia, automatizácia, optimalizácia, evolučný algoritmus

## I. ÚVOD

Keď sme vstúpili do 20. storočia, ľudská populácia bola približne 1,5 miliardy. Odkedy sme vstúpili do 21. storočia, odhaduje sa, že ľudská populácia sa zvýšila už na 7 miliárd. Teraz začíname vidieť hranice novo nájdených technológií a schopnosti našej planéty podporovať neustály rast populácie. Pôda eroduje a nasycuje sa chemickými pozostatkami. Zásoby vody zlyhávajú a sú natiiahnuté ďaleko nad jej limit. Naša atmosféra bola poškodená a klimatické podmienky sa zmenili. A aj napriek tomu ľudská populácia neustále stúpa a s rastúcou populáciou sa dostávame do situácie, kedy naša planéta už nebude schopná užiť celú ľudskú populáciu, ak to ešte vôbec dokáže. Dostali sme sa do situácie, kedy by sme mali začať zamýšľať nad znížením ľudskej populácie prostredníctvom plánovania a obmedzení, inak sa o znížovanie populácie postará príroda sama chorobami a hladom. Práve použitie najmodernejších počítačom riadených technológií v spojení s hydroponickým spôsobom pestovania rastlín má obrovský potenciál do budúcnosti, ako zabezpečiť plodiny pre súčasnú populáciu [1].

## II. PRINCÍP OPTIMALIZÁCIE EVOLUČNÝCH ALGORITMOV

Optimalizačné algoritmy sa používajú predovšetkým tam, kde riešenie problému nie je možné dosiahnuť pomocou analytickej cesty z dôvodu, že nie je vhodná alebo sa nedá vôbec použiť na riešenie úlohy. Riešenie úloh, ktoré si vyžadujú použitie optimalizačných metód, predstavuje vlastne manipuláciu s argumentmi optimalizovaných funkcií, pričom definičné obory argumentov môžu byť rôzneho druhu. Môžu sa vyskytovať napríklad obory, celočíselných, reálnych ale aj komplexných a iných čísel. V procese samotnej optimalizácie sa môžeme stretnúť aj s rôznymi obmedzeniami, ktoré sa nemusia výhradne týkať len argumentov, ale aj funkčných hodnôt optimalizovanej funkcie. Donedávna sa na riešenie týchto optimalizačných problémov používala práve analytická cesta, ktorá bola veľmi zdĺhavá a komplikovaná, a preto sa v poslednej dobe začalo pracovať na vývoji algoritmov, ktoré práve umožňujú riešiť tieto zložité problémy veľmi efektívnou a rýchlou cestou. Evolučné algoritmy (EA) sa radia do kategórie optimalizačných algoritmov, ktoré sa delia na deterministické a stochastické. Jednou z hlavných výhod EA metód v porovnaní s ostatnými metódami je, že si vystačia s malou

špecifikáciou problému a môžu byť aplikované na širokú škálu problémov [2]. EA metódy potrebujú len cieľovú (fitness) funkciu pre riešenie daného problému, ktorý má byť optimalizovaný. Ďalšie dodatočné informácie môžu byť ľahko uvedené do EA heuristiky pre zlepšenie výkonu optimalizácie. EA majú nielen zanedbateľné nároky na povahu prehl'adávacieho problému, ale dokonca môžu byť použité na riešenie zložitých problémov. Okrem toho EA reagujú na mnoho externých parametrov evolučného procesu, a preto sú veľmi vhodné na riešenie odlišných problémov bez nutnosti špeciálneho ladenia alebo znalostí experta. Vzhľadom na rozmanitosť EA metód je jednoduché si vybrať konkrétnu metódu EA, ktorá sa hodí na riešenie daného problému [3].

Ak chceme maximalizovať cieľovú funkciu, môžeme náhodne vytvoriť sadu riešení a aplikovať cieľovú funkciu na ohodnotenie. V tomto prípade čím je cieľová funkcia vyššia, tým je lepšia. Na základe ohodnotenia niektorí z lepších jedincov sú vybraní na vytvorenie nasledujúcej generácie aplikovaním rekombinácie a mutácie na týchto jedincov. Rekombinácia sa aplikuje na dvoch a viacerých vybraných jedincov, ktorí sa tiež nazývajú rodičia a výsledkom je jeden alebo viac nových jedincov nazývaných potomkovia. Mutácia sa aplikuje na jedného vybraného jedinca a výsledkom je práve jeden nový kandidát. Vykonanie rekombinácie a mutácie vedie k vytvoreniu novej množiny jedincov (potomkov), ktorí súťažia na základe fitness funkcie alebo aj veku so starými jedincami o miesto v ďalšej populácii. Tento proces sa môže opakovať, pokiaľ sa nenájde jedinec postačujúcej kvality (riešenie) alebo bol dosiahnutý vopred nastavený počet opakovaní. V tomto procese sú dva základné vplyvy, ktoré tvoria základ evolučných algoritmov:

- Variačné operátory (rekombinácia a mutácia) vytvoria potrebnú rozmanitosť v populácii, zatiaľ čo
- selekčné operátory pôsobia ako sila, ktorá podnecuje kvalitu populácie

Aplikovaním variácie a selekcie súčasne všeobecne vedie k zlepšeniu fitness hodnôt v po sebe nasledujúcich populáciách. Môžeme si všimnúť, že mnohé prvky evolučného procesu sú stochastické. V procese selekcie lepší jedinci majú väčšiu šancu byť vybratí než jedinci s horšou fitness, ale zvyčajne aj slabšie jedince majú šancu sa stať rodičmi alebo prežiť. Pri rekombinácii jedincov je proces výberu časti jedincov, ktorá bude rekombinovaná náhodná. Podobne aj pre mutáciu platí, že časti, ktoré budú mutované u jedinca a nové časti, sú vyberané náhodne [3].

### III. POPIS IMPLEMENTÁCIE ZÁKLADNÝCH KOMPONENTOV EVOLUČNÉHO ALGORITMU

#### Reprezentácia jedinca:

Reprezentáciu jedinca možno zaradiť medzi jednu z najdôležitejších a zároveň najzložitejších častí implementácie evolučného algoritmu. Reprezentácia jedinca je rozdelená do dvoch štruktúr. Pôvodná myšlienka bolo použitie štruktúry listu pre reprezentáciu jedincov, avšak vzhľadom na potrebu vlastnej implementácie takejto štruktúry v jazyku C++, bola zvolená štruktúra poľa ako základ pre reprezentáciu jedinca. V konečnom dôsledku sa ukázalo pole dokonca ešte ako vhodnejšia voľba pre implementáciu, nakoľko pri samotnej evolúcii dochádza k veľmi častému náhodnému prístupu k jedincom ako aj k ich vlastnostiam a pole nám umožňuje oveľa rýchlejšiu odozvu na takéto operácie. Jedinec teda bol reprezentovaný prostredníctvom dynamicky alokovaného trojrozmerného poľa, na základe vstupných parametrov programu. Na (Obr. 1 **Error! Reference source not found.**) je znázornená reprezentácia populácie jedincov trojrozmerným poľom.

Jedinec	Dĺžka dňa	<ol style="list-style-type: none"> <li>1. Svetlo</li> <li>2. Intenzita svetla</li> <li>3. Zavlažovanie</li> <li>4. Vysušovanie</li> <li>5. Aktuálny odber</li> <li>6. Štart/koniec zavlažovania</li> </ol>
---------	-----------	--

Obr. 1 Reprezentácia jedinca trojrozmerným poľom

Parameter *jedinec* v poli predstavuje dynamickú premennú, ktorú zadáva užívateľ pred štartom programu. Táto premenná uchováva hodnotu statického počtu jedincov v danej populácii, čo znamená, že počet jedincov v populácii sa z generácie na generáciu nemení. Ako už vyplýva z popisu vyššie, tento parameter je celočíselného typu. Parameter *dĺžka dňa* taktiež predstavuje dynamický alokovateľnú premennú, ktorá sa zadáva pred samotným spustením programu. Táto premenná slúži na špecifikáciu dĺžky dňa v minútach. Štandardne je táto hodnota inicializovaná na hodnotu 1440, čo reprezentuje počet minút v jednom dni. V takomto prípade to znamená, že je popisovaná štruktúra jedinca v minútových intervaloch. Tento parameter je možné zvýšiť alebo znížiť podľa potreby, avšak odporúča sa s týmto parametrom veľmi neexperimentovať. Posledný parameter možno nazvať *vlastnosti jedinca*, ktoré sú staticky alokované do 6 typov.

Druhá reprezentácia jedinca v populácii je definovaná prostredníctvom triedy *species*. Táto reprezentácia jedinca je inicializovaná ako jednorozmerné pole o veľkosti populácie typu triedy *species*, ktorá je znázornená na (Obr. 2). Trieda *species* disponuje vstupnými parametrami charakterizujúcimi vlastnosti jedinca, ale aj metódami pre inicializáciu týchto parametrov. Medzi hlavné vlastnosti jedinca uložené v tejto triede patria:

- začiatok sekvencie svietenia,
- dĺžka sekvencie svietenia,
- intenzita svetla,
- začiatok intervalu zavlažovania-vysušovania,
- dĺžka intervalu medzi jednotlivými fázami zavlažovania,
- parameter vyjadrujúci celkovú fitness jedinca.

Táto druhá reprezentácia má vplyv predovšetkým na inicializáciu prvotnej populácie, ktorej sa budeme venovať v nasledujúcej kapitole. Z hierarchického hľadiska môžeme nazvať túto druhú reprezentáciu za základnú, nakoľko je samostatná a nevyžaduje si žiadne údaje z prvej reprezentácie. Naopak to však povedať nemôžeme, pretože prvá reprezentácia je závislá od druhej, a teda vyžadujú sa pri jej návrhu vlastnosti jedinca uložené v tejto druhej reprezentácii.

Jedinec	<ol style="list-style-type: none"><li>1. Začiatok sekvencie svietenia</li><li>2. Dĺžka sekvencie svietenia</li><li>3. Intenzita svetla</li><li>4. Začiatok intervalu zavlažovania-vysušovania</li><li>5. Dĺžka intervalu zavlažovania-vysušovania</li><li>6. Fitness hodnota jedinca</li></ol>
---------	--

Obr. 2 Reprezentácia jedinca prostredníctvom triedy *species*

#### Vyhodnocovacia (fitness) funkcia:

Vyhodnocovacia funkcia predstavuje nezastupiteľnú úlohu v procese optimalizácie hydroponického systému. Pri návrhu fitness funkcií sme sa zamerali na 3 základné kritériá, ktoré chceme optimalizovať. Tieto kritériá optimalizácie je možné medzi sebou kombinovať, avšak nie úplne. Pred spustením programu sú kritériá nastavené užívateľom a systém tak dokáže pracovať zakaždým s inou fitness funkciou. Optimalizáciu systému teda môžeme vykonávať s ohľadom na:

- Optimalizáciu energie
- Optimalizácia časového úseku
- Optimalizácia aktuálneho odberu

#### Optimalizácia energie:

Pri hodnote tejto fitness funkcie sa vypočítava hodnota spotreby energie každého jedinca v rámci jedného dňa. Jedinca s nižšou spotrebou energie sú preferované pred jedincami s vyššou, čo vlastne znamená, že fitness funkcia predstavuje minimalizáciu. Pri vypočítavaní fitness funkcie sa berú do úvahy všetky vstupné parametre zadané užívateľom, ktoré majú vplyv na celkovú spotrebu energie jedinca. V konečnom dôsledku to znamená, že sa snažíme nájsť jedinca s čo najmenším možným intervalom svietenia, najnižšou intenzitou a najväčšími možnými intervalmi zavlažovania.



#### Optimalizácia časového úseku:

Hlavnou myšlienkou tejto fitness funkcie je optimalizácia systému pre časový úsek zadany užívateľom. Úlohou tejto fitness funkcie je v určitom časovom úseku vykonať čo najviac činností hydroponického systému. Napríklad, ak chceme z nejakého dôvodu optimalizovať systém od 10:00 do 15:00, tak v tomto intervale by sa malo určite svietiť a vykonávať čo najviac sekvencií zavlažovania systému. Pri hodnotení touto fitness funkciou sa teda berú do úvahy len činnosti hydroponického systému v danom časovom úseku. Činnosť systému mimo toho intervalu je generovaná bez vplyvu na fitness funkciu, čo znamená, že nie je pre nás už podstatná.

#### Optimalizácia aktuálneho odberu:

Táto fitness funkcia predstavuje v podstate kontrolu jedincov, či spĺňajú respektíve nespĺňajú podmienku pre maximálny odber energie v danej minúte činnosti hydroponického systému. Pri tomto kritériu nie je pridelovaná fitness hodnota z nejakého intervalu, ktorá určuje, či je riešenie lepšie alebo nie, naopak, táto fitness nadobúda len hodnoty typu *TRUE* a *FALSE*, ktorá určuje, či jedinec vyhovuje alebo nie. V prípade, ak jedinec nevyhovuje, je okamžite vymazaný z populácie a je nahradený novým jedincom. Ak je hodnota pre aktuálny odber nastavená neadekvátne k danému hydroponickému systému, to znamená, že je príliš malá, a teda neexistuje jedinec, ktorý by spĺňal takúto podmienku, tak by sme sa mohli dostať do nekonečnej slučky v procese evolúcie. Tento problém je ošetrený prerušením činnosti programu so správou o pravdepodobne zlom vyplnení tejto hodnoty.

#### **Selekčný mechanizmus a náhrada:**

Tieto dva komponenty evolučného algoritmu sú implementované súčasne. Úlohou selekčného mechanizmu je vytvorenie rodičov, zatiaľ čo náhrada sa stará o doplnenie populácie novými jedincami. Pred každým spustením selekcie prebieha zoradenie súčasnej populácie podľa aktuálnej fitness funkcie jedincov z pôvodnej populácie. Proces selekcie funguje na princípe výberu 20% jedincov z pôvodnej populácie, ktorí sa stanú rodičmi a zároveň sa stanú súčasťou novej generácie. Pre začepčenie uchovania najlepších jedincov z populácie bol implementovaný elitizmus, ktorý automaticky posúva do novej generácie 5% najlepších. Zvyšných 15% je vybraných procesom selekcie, pričom sú implementované dve možnosti selekcie:

- Stochastický výber bez náhrady
- Q-árny turnaj

#### Stochastický výber bez náhrady:

Táto selekcia predstavuje zostrojenie rulety s obvodom  $p$ , ktorý predstavuje súčet všetkých fitness funkcií jedincov v populácii. Jednotlivé výseky na rulete predstavujú pravdepodobnosť výberu jedinca za rodiča. Čím má jedinec fitness funkciu vyššiu, tým má väčšiu šancu stať sa rodičom. Stochastický výber s náhradou znamená, že po každom vybratí sa zníži pravdepodobnosť vybratia daného jedinca o 10% znížením jeho fitness funkcie. V skutočnosti sa neznižujú reálne fitness funkcie jedincov, ale len ich kópie vytvorené v procese selekcie, aby nedošlo k podceňovaniu jedincov. Ak šanca (fitness funkcia) klesne pod hodnotu 1, tak jedinec má nulovú pravdepodobnosť stať sa rodičom. Tento proces sa opakuje, pokiaľ nie je zaplnených 20% populácie.

#### Q-árny turnaj:

Pri použití tejto metódy selekcie je pri spustení programu potrebné zadať parameter  $q$  určujúci árnosť turnaja. Proces výberu jedincov prebieha nasledujúcim spôsobom. Z populácie sa vyberie náhodným spôsobom  $q$  jedincov, jedince vybrané v elitizme už nie sú zahrnuté do tohto náhodného výberu. Vybrané jedince medzi sebou súťažia a víťaz tohto turnaja sa stáva rodičom. Podobne ako pri stochastickom výbere bez náhrady aj tento proces sa opakuje, pokiaľ sa nezaplní populácia na 20%. Výhodou oproti stochastickému výberu s náhradou je, že q-árny turnaj nám prostredníctvom parametra  $q$  na vstupe umožňuje nastavovanie selekčného tlaku. Čím je parameter  $q$  väčší, tým je selekčný tlak väčší, a teda preferujeme výber lepších jedincov za rodičov.

Spojenie fázy selekcie a náhrady je z dôvodu, že rodičia sú automaticky vybratí do novej populácie a tvoria 20% jej veľkosti. Zvyšných 80% populácie je potom doplnených aplikovaním variačných operátorov na týchto rodičov.

#### **Variačné operátory:**

Kríženie sa používa na vytváranie nového jedinca, pričom sú náhodne vybrané vlastnosti z dvoch rodičov. Keďže sme na kríženie použili práve dvoch rodičov, tak takéto kríženie sa tiež

nazýva aj sexuálne kríženie. Ako bolo už spomenuté vyššie, prvých 20% populácie je naplnených v procese selekcie rodičmi. Kríženie funguje tak, že z vybraných rodičov vyberá stále dvoch náhodných, ktorí sa budú medzi sebou krížiť. Tento proces sa opakuje, pokiaľ nie je zaplnená populácia na 100% novými jedincami. Pri vyberaní náhodných rodičov je ošetrované, aby sa nemohli krížiť rovnaké jedince, nakoľko by v procese kríženia nedošlo k žiadnej zmene, a teda takéto kríženie by bolo zbytočné. Pri krížení dvoch rodičov vzniká práve jeden nový potomok. Pre operátor kríženia bolo implementovaných šesť metód, pričom na kríženie dvoch náhodne vybraných rodičov sa použije práve jedna náhodne vybraná metóda kríženia. Nakoľko niektoré vlastnosti jedinca sú na sebe závislé, nie je možné náhodné kríženie týchto vlastností.

Operátor mutácie predstavuje asexuálne kríženie, a teda je aplikovaný na náhodne vybraných jedincov z celej populácie. Pri aplikovaní mutácie je kľúčový parameter určujúci pravdepodobnosť mutácie, ktorý sa zadáva pred spustením programu užívateľom. V literatúre sa stretávame s rôznymi pravdepodobnosťami mutácie v prírode. Niektoré uvádzajú pravdepodobnosť mutácie do 1% a iné zase do 3%. V našom prípade je možné tento parameter nastavovať manuálne, čo nám umožňuje ovplyvňovať proces evolúcie. So zvyšujúcou sa hodnotou parametra mutácie sa zvyšuje šanca na nájdenie lepšieho jedinca, avšak dochádza aj k znehodnocovaniu veľkého počtu jedincov. Tak ako aj operátor kríženia a veľa ďalších komponentov evolučného algoritmu aj mutácia je charakterizovaná stochastickým výberom. Ako bolo už spomenuté, operátor mutácie sa aplikuje na celú populáciu vrátane rodičov vybraných vo fáze selekcie a nových jedincov vytvorených krížením rodičov. Ak dôjde k mutácii jedinca, tak v tomto kroku je stochastickým výberom určené, ktoré zo 4 implementovaných metód mutácie sa aplikujú na jedinca. To znamená, že v konečnom dôsledku sa môže stať, že ani jedna z metód mutácie sa neaplikuje, a teda jedinec ostane nezmenený, aj keď bol určený na mutáciu. Existuje však aj presný opak, a teda aplikujú sa všetky metódy mutácie a dôjde k tzv. hypermutácii jedinca. Tieto dva extrémny však majú veľmi malú šancu na realizovanie.

#### IV. EXPERIMENT OPTIMALIZÁCIE HYDROPONICKÉHO SYSTÉMU

Tento experiment je zameraný na návrh plánu činnosti hydroponického systému na Technickej univerzite v Košiciach. Parametre charakterizujúce ohraničenia systému sú uvedené na (Obr. 3). Pri procese optimalizácie sme sa zamerali na optimalizáciu časového úseku a optimalizáciu aktuálneho odberu energie.

##### A. Nastavenie parametrov popisujúcich hydroponický systém

Prvým krokom pri experimente bolo nastavenie parametrov popisujúcich vlastnosti hydroponického systému na Technickej univerzite v Košiciach pre potrebu optimalizácie. Parametre boli nastavené podľa (Obr. 3).

##### B. Nastavenie parametrov evolúcie

Po zadaní vstupných parametrov do programu sme experimentovali s parametrami evolúcie. Použili sme viacero rôznych nastavení parametrov a ako najúspešnejšie riešenie sa nám ukázalo použitie hodnôt opisujúcich evolúciu uvedených na (Obr. 4).

##### C. Výstup z aplikácie

Vykonávali sme experimenty s väčšími počtami jedincov a generácií, použitím stochastického výberu bez náhrady, modifikovali sme aj parameter určujúci árnosť turnajov a pravdepodobnosť mutácie, avšak výsledok sa už nezlepšoval. Na obrázkoch (Obr. 5 **Error! Reference source not found.**) a (Obr. 6) je znázornený grafický výstup nájdeného riešenia a tabuľka činnosti hydroponického systému vyexportovaná do Excelu.

**Example 1:**

Basic options:

- Critical power (W/h) – 10 w

Light options:

- Number of lights – 4
- Light interval Min (min) – 8 h (480 min)
- Light interval Max (min) – 12 h (720 min)
- Light intensity Min (lux) – 10000
- Light intensity Max (lux) – 120000
- Consumption of light Min (W/h) – 5
- Consumption of light Max (W/h) – 10

Pump options:

- Number of pumps – 1
- Flooding time (min) – 3
- Draining time (min) – 10
- Flood-drain interval Min (min) – 30
- Flood-drain interval Max (min) – 120
- Consumption of pump (W/h) – 25watt max

Evolution options:

- Optimize for TOD (hour) – from 08:00 to 19:00
- Optimize for PEAK (W/m) – 60watt (1w/m)

Obr. 3 Vstupné parametre pre optimalizáciu hydroponického systému

Basic options | Light options | Pump options | Evo options

Evolution options:

Species: 200

Generations: 100

Selection: Tournaments: 8, Roulette:

Mutation: 5

Day frequency: 1440

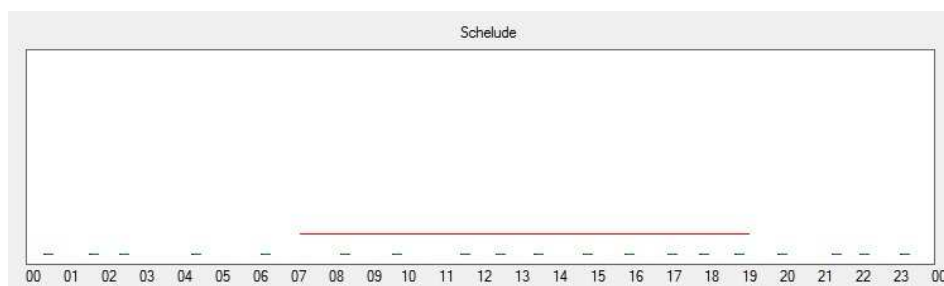
Initialization:

optimize for ENERGY

optimize for TOD from 08:00 to 19:00

optimize for PEAK 01,00

Obr. 4 Nastavanie parametrov evolúcie



Obr. 5 Grafický rozvrh činnosti hydroponického systému

Light options:			
Light intensity:	10079lux		
Light duration	Start at:	End at:	
	7hod 11min	19hod 2min	
Pump options:			
	Start flood at:	End flood at:	End drain at:
Interval-1:	0hod 26min	0hod 29min	0hod 39min
Interval-2:	1hod 38min	1hod 41min	1hod 51min
Interval-3:	2hod 25min	2hod 28min	2hod 38min
Interval-4:	4hod 20min	4hod 23min	4hod 33min
Interval-5:	6hod 9min	6hod 12min	6hod 22min
Interval-6:	8hod 15min	8hod 18min	8hod 28min
Interval-7:	9hod 37min	9hod 40min	9hod 50min
Interval-8:	11hod 25min	11hod 28min	11hod 38min
Interval-9:	12hod 21min	12hod 24min	12hod 34min
Interval-10:	13hod 22min	13hod 25min	13hod 35min
Interval-11:	14hod 39min	14hod 42min	14hod 52min
Interval-12:	15hod 45min	15hod 48min	15hod 58min
Interval-13:	16hod 53min	16hod 56min	17hod 6min
Interval-14:	17hod 44min	17hod 47min	17hod 57min
Interval-15:	18hod 40min	18hod 43min	18hod 53min
Interval-16:	19hod 48min	19hod 51min	20hod 1min
Interval-17:	21hod 13min	21hod 16min	21hod 26min
Interval-18:	21hod 57min	22hod 0min	22hod 10min
Interval-19:	23hod 2min	23hod 5min	23hod 15min

Obr. 6 Rozvrh činnosti hydroponického systému pomocou Excel tabuľky

#### D. Vyhodnotenie experimentu

Experiment môžeme zhodnotiť ako úspešný, nakoľko nájdené riešenie spĺňa všetky požiadavky hydroponického systému a takisto aj požiadavky zadané užívateľom týkajúce sa optimalizácie. Nájdené riešenie je použiteľné pri prevádzke hydroponického systému a predstavuje značné urýchlenie času pri rozvrhovaní a optimalizácii činnosti hydroponického systému na Technickej univerzite v Košiciach. Nevylučuje sa nájdenie lepšieho výsledku pri vhodnejšom nastavení parametrov evolúcie. Nakoľko nevieme určiť hodnotu najlepšieho jedinca, je veľmi ťažké presne určiť, aké úspešné je navrhnuté riešenie [4].

#### V. ZÁVER

Cieľom tejto práce bolo analyzovať možnosti pestovania rastlín automatizovanými hydroponickými systémami a evolučné algoritmy ako možnosť pre optimalizáciu týchto systémov. Pokúsili sme sa navrhnúť program pre optimalizáciu hydroponického systému na Technickej univerzite v Košiciach.

V tomto článku sme sa teda hlavne venovali možnostiam použitia evolučného algoritmu a jeho následnej implementácii pre optimalizáciu hydroponického systému na Technickej univerzite v Košiciach. Program bol navrhnutý v jazyku C++. V článku je popísaná implementácia základných komponentov evolučného algoritmu. V procese testovania boli postupne odstraňované chyby, ktoré vznikali pri procese evolúcie.

Program predstavuje účinný nástroj na optimalizáciu hydroponických systémov. V ďalšom pokračovaní by sa dal program rozšíriť o ďalšie vlastnosti systému ovplyvňujúce rast rastliny, ako je teplota okolia, vlhkosť, kvalita živného roztoku a mnoho ďalších.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] Anglická verzia stránky Wikipedia venovaná téme Agriculture [online]. [cit. 2014-05-02]. Dostupné na internete: <http://en.wikipedia.org/wiki/Agriculture>.
- [2] Slovenský článok zaoberajúci sa problematikou evolučných a genetických algoritmov. [online]. [cit. 2014-05-02]. Dostupné na internete: <http://www.posterus.sk/?p=3364>.
- [3] SZÖLLŐSI, T. Evoluční algoritmy. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 74 s.
- [4] Viktor Šoltés, *Použitie evolučného algoritmu pre optimalizáciu energie v skleníku*, Diplomová práca, Technická univerzita, Košice, 2014, 80s.

# Meranie účinnosti tienenia a odrazu elektromagnetického poľa pre stavebný objekt polystyrén

<sup>1</sup>Marek PAVLÍK, <sup>2</sup>Iraida KOLCUNOVÁ, <sup>3</sup>Michal KOLCUN, <sup>4</sup>Ján  
ZBOJOVSKÝ, <sup>5</sup>Peter KURIMSKÝ

<sup>1,2,3,4,5</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
Univerzita v Košiciach, Slovenská Republika

[marek.pavlik@tuke.sk](mailto:marek.pavlik@tuke.sk), [iraida.kolcunova@tuke.sk](mailto:iraida.kolcunova@tuke.sk), [michal.kolcun@tuke.sk](mailto:michal.kolcun@tuke.sk),  
[jan.zbojovsky@tuke.sk](mailto:jan.zbojovsky@tuke.sk), [peter.kurimsky@student.tuke.sk](mailto:peter.kurimsky@student.tuke.sk)

**Abstrakt** — Tento článok pojednáva o meraní účinnosti tienenia SE a odrazu R elektromagnetického poľa. Z výsledkov merania bola vypočítaná absorpcia elektromagnetického poľa. Meranie prebiehalo v bezodrazovej komore vo frekvenčnom rozsahu od 1 GHz do 9 GHz s krokom 0,1 GHz. Meraným objektom bol stavebný materiál polystyrén Styrodur s hrúbkou 50 mm.

**Kľúčové slová** — účinnosť tienenia, odraz elektromagnetického poľa, absorpcia elektromagnetického poľa

## I. ÚVOD

Moderný človek si rád dopraje komfort. Pod týmto pojmom je možné si predstaviť používanie rozličných zariadení uľahčujúcich život, či sa jedná o diaľkové ovládanie, využívanie mobilov, bezdrôtového internetu, prípadne príjem rozhlasu a televízie. Cestu k zaobstaraniu si týchto zariadení si nachádza čoraz viac ľudí. Tieto rôzne výdobytky modernej doby predstavujú elektrické zariadenia, ktoré vyžarujú do okolia elektromagnetické polia (vlnenia). V súčasnosti sa už tieto polia nachádzajú všade okolo nás.

Pôsobenie elektromagnetického poľa na živé organizmy je intenzívne skúmanou problematikou. V posledných rokoch bolo realizovaných množstvo štúdií, ktoré sa snažili dokázať, resp. vyvrátiť súvislosť pri zmenách zdravotného stavu v dôsledku pôsobenia elektromagnetického poľa. Zatiaľ neboli vplyvy a následky pôsobenia takéhoto poľa jednoznačne dokázané. Na pomenovanie nežiaducich polí v pracovnom prostredí sa zaviedol pojem elektrosmog. Účinky elektromagnetického žiarenia nie sú rovnaké pre celý rozsah frekvencií, môžu mať dvojaký účinok. Na jednej strane pozitívny, využívaný v medicíne, na druhej strane negatívny. Negatívny má dvojakú formu účinkov, tepelných a netepelných. Tepelné účinky rastú s nárastom frekvencie a zahrievajú organizmus. Naopak netepelné majú nepriaznivý účinok na stav centrálnej nervovej sústavy. Pre tieto negatívne účinky Ministerstvo zdravotníctva určilo limitné hodnoty expozície pri dlhodobom pôsobení na človeka. Limitné hodnoty expozície sa nachádzajú vo vyhláske 534/2007 Ministerstva zdravotníctva [1][2].

Pôsobenie elektromagnetického poľa na živé organizmy je intenzívne skúmanou problematikou. V posledných dvoch desaťročiach bolo realizovaných množstvo štúdií, ktoré sa snažili dokázať, resp. vyvrátiť súvislosť pri zmenách zdravotného stavu v dôsledku pôsobenia elektromagnetického poľa. Zatiaľ neboli vplyvy a následky pôsobenia takéhoto poľa jednoznačne dokázané. Moderný človek je denne vystavovaný elektromagnetickým poliám rôznych intenzít z rôznych zdrojov, preto dôkaz o negatívnych účinkov týchto polí (rakovina, psychické poruchy, poškodenie orgánov, atď.) by malo dramatický dopad na neho [1][2].

Otázkou návrhu najvyšších prípustných hodnôt pre elektromagnetické pole sa začali zaoberať: Svetová zdravotnícka organizácia (WHO, World Health Organisation) a Medzinárodná komisia pre ochranu pred neionizujúcim žiarením (ICNIRP, International Commission on Non- Ionizing Radiation Protection), ale až po množstve negatívnych ohlasov,

ktoré začali stúpať. Boli to ohlasy ľudí, ktorí pracovali s výpočtovou technikou. Tí sa sťažovali na zhoršený zdravotný stav, najmä pri práci s počítačom. Po tomto organizácie stanovili určité limity, ktoré elektrické zariadenia musia spĺňať, taktiež musia byť elektromagneticky kompatibilné, aby mohli byť používané v praxi bez ohrozenia zdravia človeka .

Elektromagnetické pole je biologicky aktívne, nakoľko má schopnosť ovplyvňovať živý organizmus v celom svojom frekvenčnom spektre. Účinok elektromagnetického žiarenia však nie je rovnaký pre celý rozsah frekvencií, ktoré sa vyskytujú od najnižších kmitočtov až do kmitočtu 1024 Hz [1][2].

## II. ÚČINKY ELEKTROMAGNETICKÉHO POĽA

Vo všeobecnosti je možné rozlíšiť dva typy účinkov elektromagnetického poľa. Tepelné a netepelné. Tepelné účinky na organizmus sú späté s frekvenciou. S nárastom frekvencie rastú aj tepelné účinky. Obsah vody v tkanive, veľkosť povrchu, ktorý je vystavený žiareniu a taktiež aj intenzita poľa majú značný vplyv na množstvo absorbovanej energie v tkanive. Problém nastáva, ak nedokáže organizmus odvieť do okolitého prostredia množstvo privádzaného tepla, vtedy môže dôjsť k tepelnému preťaženiu niektorých orgánov. Najviac citlivé orgány na zvýšenie teploty sú: očná šošovka, mozog a semenníky. Napr. očná šošovka je zvlášť citlivá na vysokofrekvenčné žiarenie. Jej vystavenie relatívne malej záťaži môže spôsobiť jej zákal [3][4].

Netepelné účinky sú prevažne určené okamžitou amplitúdou vysokofrekvenčného žiarenia. Opakované ožarovanie pomerne nízkymi intenzitami má za následok, že ich účinok stúpa, hlavne pri vystavení sa pulznému poľu. Celkový vyžiarený výkon tohto poľa je pomerne malý, ale jeho okamžitá amplitúda je veľká. Vnímavosť poľa u jednotlivcov sa výrazne odlišuje [4].

Vplyv dlhodobého ožarovania EMP s malou hustotou výkonu sa prejaví predovšetkým na stave centrálnnej nervovej sústavy. Tieto zmeny spôsobujú telesnú slabosť ako napr.: vyčerpanosť, ľahostajnosť, zvýšenú únavu, poruchy spánku, pokles pozornosti, zhoršenie pamäti, bolesti hlavy, precitlivenosť, znížená potencia [3][4][5].

## III. TIENENIE ELEKTROMAGNETICKÉHO POĽA

Elektromagnetické tienenie je jedno z dôležitých prostriedkov, ktoré bráni prenikaniu elektromagnetického poľa. Umožňuje zmenšenie rušivého vyžarovania na strane zdrojov rušivých signálov, taktiež zvyšuje elektromagnetickú odolnosť na strane prijímačov rušivých signálov. Tienenie sa používa k odrušovaniu poľa vo vymedzenom priestore. Prostriedky, ktorými je možné to dosiahnuť, nazývame tieniace kryty, resp. tienenie. Tienenie sa používa ako pre ochranu jednotlivých súčiastok, tak aj celých elektronických zariadení, ktoré môžu byť súčasne zdrojom i prijímačom elektromagnetického rušenia [1]

Elektromagnetické tienenie možno charakterizovať tzv. koeficientom tienenia  $K_s$ . Je to pomer intenzity elektrického poľa  $E_2$  (alebo magnetického poľa  $H_2$ ) v určitom bode tienenia priestoru k intenzite  $E_1$  ( $H_1$ ) poľa dopadajúceho na tieniacu priechku. Lepším ukazovateľom tienenia je účinnosť tienenia  $SE$ . Účinnosť tienenia je možné charakterizovať ako súčet odrazu elektromagnetického poľa  $R$  a absorpcie elektromagnetického poľa  $A$ :

$$SE = A + R \quad (1)$$

Koeficient absorpcie sa určí podľa vzťahu (2) ako:

$$A = 8,69 \frac{t}{\delta} \quad (2)$$

kde  $t$  je hrúbka tieniaceho materiálu a  $\delta$  je hĺbka vniku elektromagnetického poľa do materiálu. Hĺbka vniku sa vypočíta podľa vzťahu (3) ako:

$$\delta = \sqrt{\frac{2}{\omega \cdot \mu \cdot \sigma}} \quad (3)$$

kde  $\mu$  je permeabilita, v ktorej je zahrnutá aj permeabilita tieniaceho materiálu,  $\sigma$  je merná vodivosť tieniaceho materiálu.

Koeficient odrazu  $R$  sa určí podľa vzťahu (4) ako:



$$R = 20 \cdot \log \left( \frac{1}{4} \sqrt{\frac{\sigma}{\omega \cdot \mu_r \cdot \epsilon_0}} \right) \quad (4)$$

kde  $\mu_r$  je permeabilita tieniaceho materiálu,  $\epsilon_0$  je permitivita vákua.

V prípade, že hodnota vysielaného signálu je nastavená v logaritmickej jednotkách, tak účinnosť tienenia  $SE$  sa určí podľa vzťahu [6]:

$$SE = |E_1| - |E_2| \text{ (dB)} \quad (5)$$

$$SE = |H_1| - |H_2| \text{ (dB)} \quad (6)$$

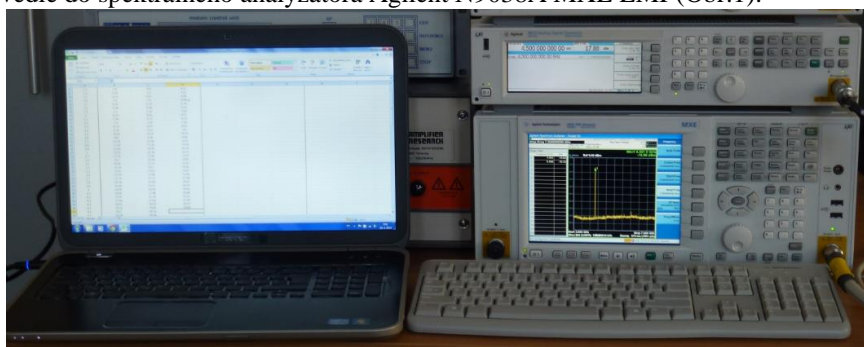
$$SE = |V_1| - |V_2| \text{ (dB)} \quad (7)$$

$$SE = P_1 - P_2 \text{ (dB)} \quad (8)$$

pričom  $V_2$  je napätie vlny elektromagnetického poľa v určitom bode tieneného priestoru,  $V_1$  je napätie vlny elektromagnetického poľa dopadajúceho na tieniacu bariéru alebo stenu,  $P_2$  je výkon elektromagnetického poľa v určitom bode tieneného priestoru,  $P_1$  je výkon elektromagnetického poľa dopadajúceho na tieniacu bariéru alebo stenu [6].

#### IV. EXPERIMENT

Tento príspevok popisuje meranie účinnosti tienenia elektromagnetického poľa a odrazu elektromagnetického poľa. Následne z týchto meraní bola vypočítaná absorpcia elektromagnetického poľa vyjadrením zo vzťahu (1). Merania prebiehali v bezodrazovej komore, ktorá pozostávala z z generátora impulzov Agilent N5181A (Obr.1), ktorým sa pripája do vysielacej antény RF spin DRH18-E (Obr.2). Medzi anténami sa nachádza merané tienenie, ktorým bol polystyrén & styrodur s hrúbkou 50 mm. Výstup z prijímacej antény R&S HF907 (Obr.2) vedie do spektrálneho analyzátoru Agilent N9038A MXE EMI (Obr.1).



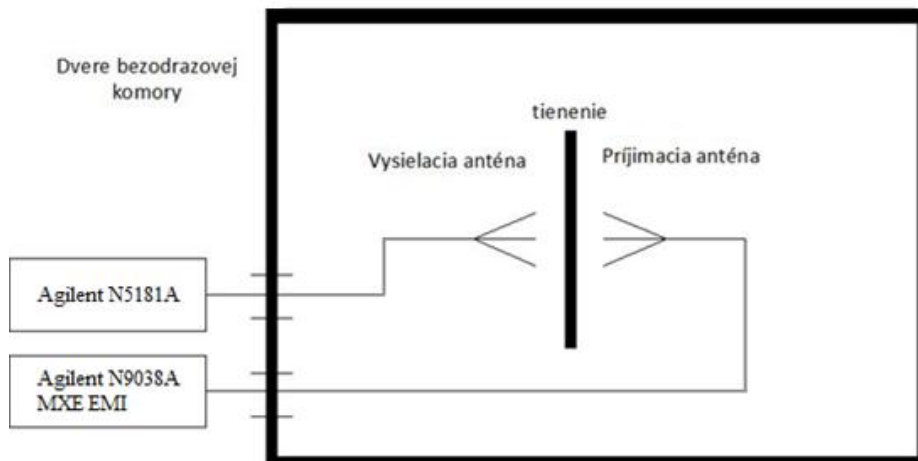
Obr. 1 Pohľad na pracovisko - Agilent N9038A MXE EMI (hore) a Agilent N5181A (dole)



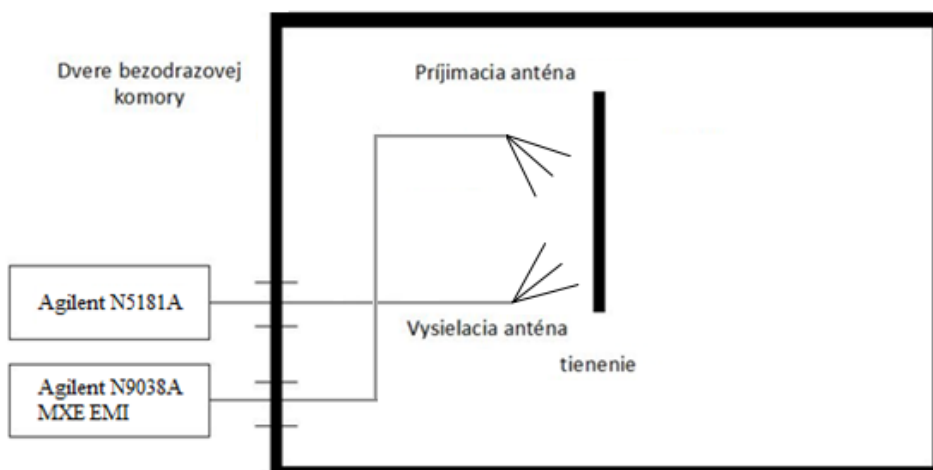
Obr. 2 Pohľad na pracovisko - RF spin DRH18-E (hore) a R&S HF907 (dole)



Blokové schémy pre účely merania účinnosti tienenia a odrazu elektromagnetického poľa sú zobrazené na Obr.3 a Obr.4.



Obr. 3 Bloková schéma pre účely merania účinnosti tienenia elektromagnetického poľa

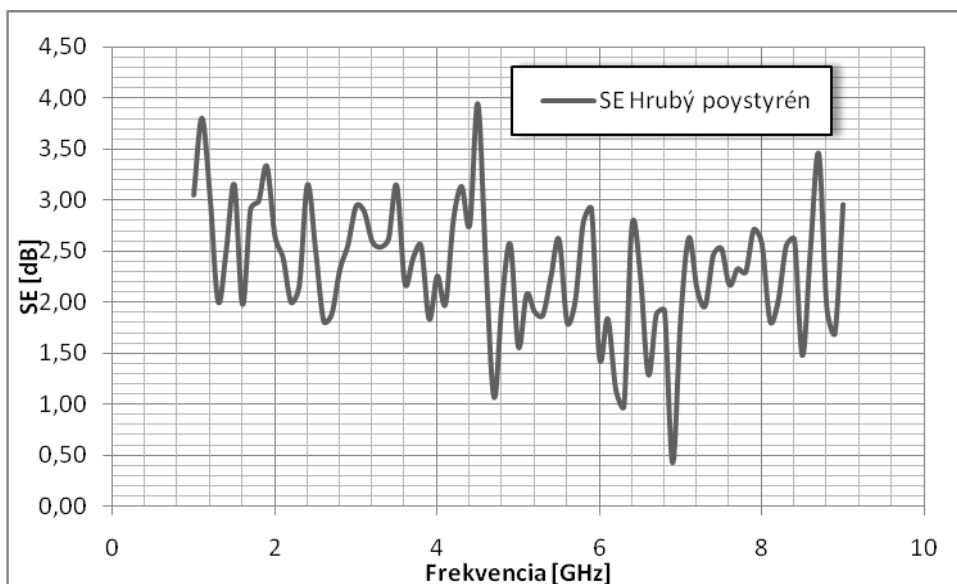


Obr. 4 Bloková schéma pre účely merania odrazu elektromagnetického poľa

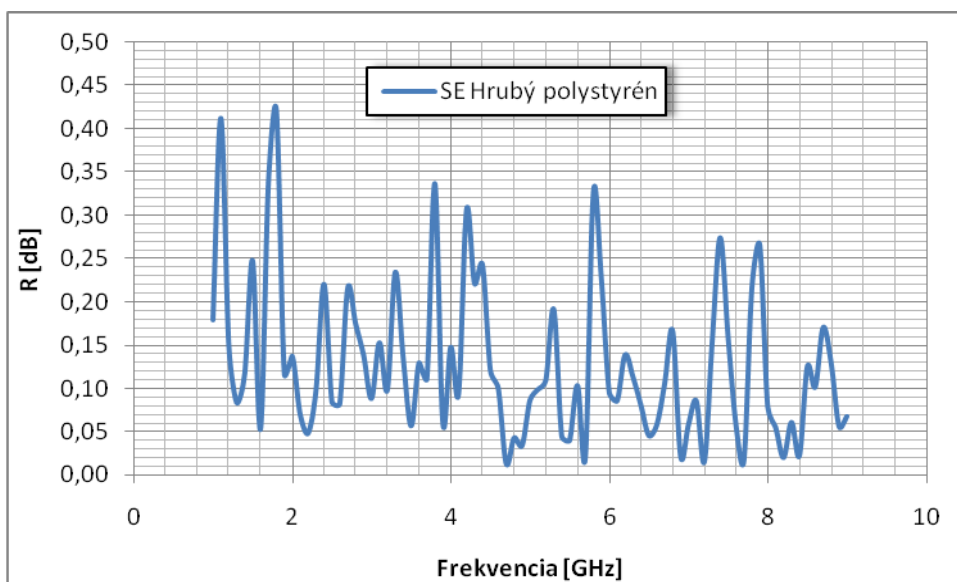
## V. VÝSLEDKY

Meranie účinnosti tienenia  $SE$  a odrazu  $R$  elektromagnetického poľa prebiehalo vo frekvenčnom rozsahu od 1 GHz do 9 GHz, s krokom 0,1 GHz. Následne na základe meraní účinnosti tienenia a odrazu elektromagnetického poľa bola vypočítaná absorpcia elektromagnetického poľa.

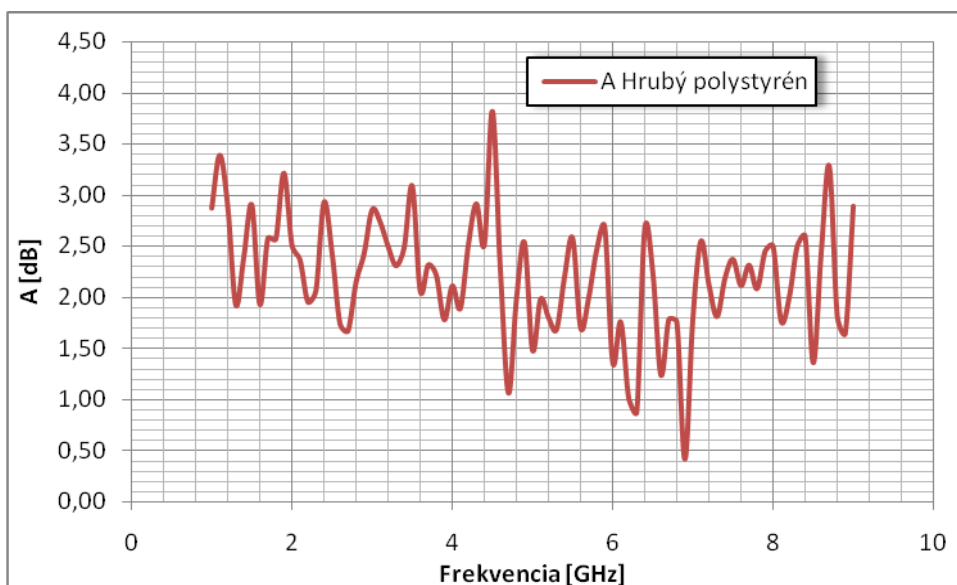
Na Obr. 5 je možné vidieť frekvenčnú závislosť účinnosti tienenia elektromagnetického poľa vo frekvenčnom rozsahu od 1 GHz do 9 GHz. Na Obr. 6 je frekvenčná závislosť odrazu elektromagnetického poľa  $R$  vo frekvenčnom rozsahu od 1 GHz do 9 GHz a na Obr.7 je frekvenčná závislosť absorpcie elektromagnetického poľa  $A$  vo frekvenčnom rozsahu od 1 GHz do 9 GHz. Z nameraných kriviek je možné vidieť, že účinnosť tienenia elektromagnetického poľa pre hrubý polystyrén sa pohybuje približne od 1 dB do 4 dB. Namerané hodnoty odrazu elektromagnetického poľa sa pohybujú od 0,03 dB do 0,4 dB. Hodnoty účinnosti tienenia a odrazu elektromagnetického poľa sa v celom frekvenčnom rozsahu neustále pohybovali v rozmedzí od svojej maximálnej po minimálnu hodnotu. Z výsledkom merania je zrejmé, že väčšiu časť z účinnosti tienenia tvorí absorpcia elektromagnetického poľa.



Obr. 5 Účinnosť tienenia elektromagnetického poľa SE pre frekvenčný rozsah od 1 GHz do 9 GHz



Obr. 6 Odraz elektromagnetického poľa R pre frekvenčný rozsah od 1 GHz do 9 GHz



Obr. 7 Absorpcia elektromagnetického poľa A pre frekvenčný rozsah od 1 GHz do 9 GHz

## VI. ZÁVER

V súčasnosti narastá počet zdrojov vyžarujúcich elektromagnetické pole a preto sa čoraz viac upriamuje pozornosť na ochranu pred elektromagnetickým žiarením. Tento článok je zameraný na meranie účinnosti tienenia elektromagnetického poľa stavebného objektu, ktorá je ukazovateľom kvality tieniaceho materiálu. Ďalej tento článok pojednáva o meraní odrazu elektromagnetického poľa a o výpočte absorpcie elektromagnetického poľa. Výsledkom merania sú frekvenčné závislosti účinnosti tienenia elektromagnetického poľa  $SE$ , odrazu elektromagnetického poľa  $R$  a absorpcie elektromagnetického poľa  $A$  pre stavebný objekt hrubý polystyrén Styrodur s hrúbkou 50 mm.

## POĎAKOVANIE

Táto práca vznikla vďaka podpore v rámci operačného programu Výskum a vývoj pre projekt: Ochrana obyvateľstva Slovenskej republiky pred účinkami elektromagnetického poľa, s kódom ITMS: 26220220145, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.



Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ.

## REFERENCES

- [1] J. Svačina, "Elektromagnetická kompatibilita, princípy a metódy", Brno 2001, ISBN 80-214-1873-7.
- [2] Vyhláška Ministerstva zdravotníctva Slovenskej republiky 534/2007 zo 16. augusta 2007 o podrobnostiach o požiadavkách na zdroje elektromagnetického žiarenia a na limity expozície obyvateľov elektromagnetickému žiareniu.
- [3] R. Drahoš, "Interakcia zdrojov elektromagnetických polí v mimopracovnom prostredí", D2R engineering, 5s.
- [4] Z. Cabanová, "Biologické účinky elektromagnetického poľa", Advances in Electrical and Electronic Engineering, Elektrotechnická fakulta Žilinská univerzita v Žiline, s. 24-29.
- [5] O. Osina, "Mobilné telefóny, aktuálny problém súčasnosti", In Nové poznatky v oblasti medicínskych vied a ošetrovateľstva, Fakulta zdravotníctva Katolíckej univerzity, Ružomberok, 2006, 34 s., ISBN 80-8084-125-X.
- [6] IEEE 299-2006 Standard, Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures, EMC Society, New York 2006, p.39.

# Analýza procesných dát z pohľadu organizačnej perspektívy

<sup>1</sup>Štefánia ŠUPALOVÁ, <sup>2</sup>Cecília HAVRILOVÁ, <sup>3</sup>Ján PARALIČ

<sup>1, 2, 3</sup> Katedra kybernetiky a umelej inteligencie, Faculta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>stefania.supalova@student.tuke.sk, <sup>2</sup>cecilia.havrilova@tuke.sk, <sup>3</sup>jan.paralic@tuke.sk

**Abstrakt** — Táto práca sa zameriava na riešenie vybraného problému dolovania procesov v záznamoch udalostí. Jej hlavným cieľom je skúmať vzťahy medzi jednotlivými organizačnými jednotkami, ktorými sú nemocničné oddelenia a analyzovať mieru spolupráce medzi týmito oddeleniami. Ďalším analyzovaným atribútom je vek pacientov, dôležitý ako pri odhalení vyťaženia oddelení, tak aj pri prevencii samotných ochorení. Z tohto dôvodu sa zameriame aj na analýzu tohto ukazovateľa. Našou ambíciou je získať výsledky, ktoré napomôžu rôznym zdravotníckym zariadeniam zameraným na gynekologické ochorenia v oblasti plánovania a zabezpečenia oddelení, ako aj pri samotnej prevencii ochorení.

**Kľúčové slová** — dolovanie procesov, organizačná perspektíva, sociálna sieť, záznamy udalostí

## I. ÚVOD

Dolovanie procesov v záznamoch udalostí je vedecká disciplína založená na dolovaní v dátach, procesnom modelovaní a analýze dát. Táto disciplína je založená na objavovaní, monitorovaní a zlepšovaní procesov z reálneho sveta, ktoré sú extrahované zo záznamov udalostí [1]. V praxi nachádza dolovanie procesov svoje uplatnenie v rôznych oblastiach, ako sú ekonomika, priemysel či v neposlednom rade zdravotníctvo. Práve zdravotníctvo je oblasť, ktorej sa budeme v tejto práci venovať. Najväčšia pozornosť je zameraná na analýzu spolupráce a odhaľovania vzťahov a väzieb medzi organizačnými zložkami, ktorými sú v našom prípade nemocničné oddelenia.

V súčasnej dobe sú veľmi frekventovanou otázkou ochorenia súvisiace s rakovinou. V tejto práci preto venujeme pozornosť ochoreniam súvisiacim so ženskými pohlavnými orgánmi. Odhalíme a analyzujeme vplyv ochorenia konkrétneho orgánu na vyťaženosť a vzájomnú spoluprácu oddelení. Pomocou získaných výsledkov tak môžeme pomôcť gynekologickým klinikám pri zabezpečení a rozmiestnení ich oddelení na základe vzájomnej spolupráce s inými oddeleniami. V práci budú použité dáta z Holandskej akademickej nemocnice (DAH).

## II. EXISTUJÚCE RIEŠENIA

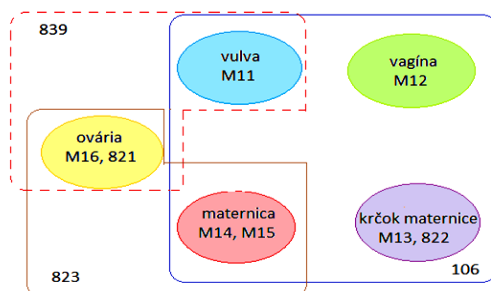
Záznamy udalostí použité v tejto práci boli podrobené niekoľkým analýzám v rámci konferencie BPI Challenge 2011<sup>1</sup>. Tieto analýzy boli popísané v nasledujúcich prípadových štúdiách. Jagadeesh a kol. [2] sa vo svojej analýze zamerali na popis jednotlivých diagnóz. Po odfiltrovaní nepotrebných dát údaje rozdelili na urgentné a neurgentné. Takto sledovali, ktoré procesy sú často sa opakujúce a pomerne jednoduché. Okrem toho vizualizovali postupnosti vzťahov medzi jednotlivými oddeleniami, ktorými pacientka s danou diagnózou prešla. V ďalšej prípadovej štúdii sa Caron a kol. [3] zamerali na analýzu oddelení na základe podprocesov, pričom sa túto analýzu rozhodli uskutočniť v rámci oddelenia Rádioterapie. Vytvorili procesný model a analyzovali organizačnú štruktúru holandského zdravotného systému. V ďalšom kroku vytvorili sociálnu sieť, ktorá slúžila na odhalenie vzťahov oddelenia Rádioterapie s ostatnými oddeleniami. Varvaressos a kol. sa vo svojej analýze [4] zamerali na vytvorenie ontológie zdravotného systému.

Naším cieľom v tejto práci bude analyzovať záznamy udalostí z pohľadu organizačnej perspektívy, pričom budeme skúmať vzťahy a väzby medzi jednotlivými organizačnými jednotkami - oddeleniami v rámci gynekologickej polikliniky. Tieto analýzy budú uskutočnené na záznamoch rozdelených podľa chorôb definovaných podľa stanovených diagnóz. Do analýz zahrnieme aj vplyv veku, ktorý je jedným z dôležitých atribútov pri prevencii ochorení.

<sup>1</sup> <http://www.win.tue.nl/bpi/2011/challenge>

### III. ANALÝZA ZÁZNAMOV UDALOSTÍ

Záznamy udalostí poskytnuté Holandskou akademickou nemocnicou (DAH) obsahujú 150 291 udalostí vo vyše 1 143 prípadoch z oddelenia gynekológie. Dáta obsahujú údaje o celkovo 624 medicínskych aktivitách. Samotný proces pojednáva o diagnostikovaní ochorenia a následnej liečbe pacientky na vybraných oddeleniach DAH. Pacientka môže prejsť počas návštevy nemocnice viacerými fázami, od diagnostiky ochorenia až po samotnú liečbu, preto sa niektoré atribúty môžu opakovať viackrát [5]. Každý prípad obsahuje niekoľko atribútov, ktoré poskytujú informácie o diagnóze chorých pacientiek. Tieto atribúty sa môžu zadeliť a klasifikovať do dvoch kategórií: *Diagnóza*, *Kód diagnózy*. Každý prípad môže obsahovať až 16 rôznych atribútov každého typu. Spomedzi všetkých prípadov nájdeme aspoň jeden z týchto atribútov v 1 136 prípadoch. Kategória *Diagnóza* obsahuje popis danej diagnózy a v kategórii *Kód diagnózy* je priradený kód danej diagnózy. Musí byť pritom dodržaná súvislosť, vzťah medzi týmito dvoma kategóriami, tzn. kategórii *Kód diagnózy* prináleží kategória *Diagnóza*, kategórii *Kód diagnózy 1* prináleží kategória *Diagnóza 1*, atď. Záznam udalostí popisuje liečebné procedúry k celkovo jedenástim odlišným diagnostickým kódom (Obr. 1).



Obr. 1 Vennov diagram znázorňujúci prienik diagnóz zo záznamov udalostí

Jednotlivé diagnózy, na ktoré sa zameriame v tejto práci počas analýzy spomínaných záznamov udalostí sú popísané v tabuľke Tab. 1, kde je taktiež uvedený počet prípadov s danou diagnózou.

Tab. 1 Popis niektorých diagnóz zo záznamov udalostí

Kód diagn.	Diagnóza	Počet
M11	Súvisí s rakovinou vulvy, popisuje diagnózy: spinocelulárny karcinóm, zhubné nádory a melanóm, hraničné malignity	4 962
M12	Súvisí s rakovinou vagíny, popisuje diagnózy: spinocelulárny karcinóm, zhubné nádory, adenokarcinóm	1 219
M13	Súvisí s rakovinou krčka maternice, popisuje diagnózy: zhubné nádory, spinocelulárny karcinóm, adenokarcinóm, hraničné malignity, sarkóm	23 648
M14	Súvisí s rakovinou tela maternice, popisuje diagnózy: adenokarcinóm vo viacerých fázach, zhubné novotvary a endometrium, karcinóm vo fáze 1B a 3B, hraničné malignity. Pri tejto diagnóze sa vyskytujú aj niektoré metastázy.	6 076
M15	Súvisí s rakovinou tela maternice, konkrétne typu sarkoma	2 016
M16	Súvisí s rakovinou vaječníkov, popisuje diagnózy: adenokarcinóm štyroch typov – seróza, endometroid, mucinóus, neepitálny zhubný nádor, ďalej aj novotvary, hraničné malignity, karcinómy a niektoré prípady metastázy	14 013
821	Súvisí s rakovinou vaječníkov, popisuje diagnózy: adencarcinoma typu serous a mucinóus, neepitálny zhubný nádor, neoplazma	13 573
822	Súvisí s rakovinou krčka maternice, popisuje diagnózy: spinocelulárny karcinóm, adenokarcinóm, hraničné malignity, zhubné novotvary	4 956
106	Súvisí s heterogénnou zmesou karcinómov, popisuje diagnózy: cervix uteri typu spinocelulárneho karcinómu, zhubné nádory, hraničné malignity, karcinóm skvamóznych buniek vulvy, malígný melanóm, adenokarcinóm krčka, zhubné novotvary, rakovina pošvy, sliznice maternice a vajčkovodov	6 150
823	Popisuje heterogénnu zmes karcinómov, popisuje diagnózy: adenokarcinóm krčka, malígný melanóm, sarcoma, adenokarcinóm vaječníkov, endometrium	1 647
839	Popisuje heterogénnu zmes prípadov súvisiacich s rakovinou ovárií, maternice, vulvy	5 381

#### IV. REALIZOVANÉ EXPERIMENTY

Na realizáciu experimentov sme vytvorili aplikáciu v programovacom jazyku MATLAB. Táto aplikácia má slúžiť pre hlbšiu analýzu dátového súboru záznamov udalostí z prostredia DAH. Ide hlavne o analýzu ktorá ponúkne používateľovi prehľad o vzťahoch a väzbách medzi prechodmi z jedného oddelenia na druhé vzhľadom na ochorenia, ktorými trpia pacienti. Vďaka tejto analýze bude môcť lekár, resp. pracovník nemocnice posúdiť spoluprácu svojho oddelenia s ostatnými oddeleniami. Tieto údaje môžu pomôcť aj manažerom pri plánovaní umiestnenia jednotlivých oddelení v rámci budovania, resp. reorganizácie kliniky. Aplikácia je preto predovšetkým určená analytikom z nemocničného prostredia, lekárom, manažerom gynekologických kliník, či pracovníkom jednotlivých oddelení.

##### A. Požiadavky na riešenie

Aplikácia je navrhovaná v nástroji Matlab, vo verzii 7.9.0 (R2009b) na operačnom systéme Microsoft Windows 7. Celkovo zahŕňa 12 funkcionalít:

- Načítanie súborov pre základné štatistiky,
- Štatistiky pre ochorenia maternice,
- Štatistiky pre ochorenia vaječníkov,
- Štatistiky pre ochorenia vonkajších pohl.orgánov,
- Štatistiky pre ochorenia heterogénnych zmesí,
- Porovnanie návštev oddelení podľa chorôb,
- Štatistiky pre počty prechodov z oddelenia,
- Pravdepodobnosti prechodov medzi oddeleniami,
- Porovnanie vekovej štruktúry pre zvolené ochorenia.

##### B. Načítanie súboru

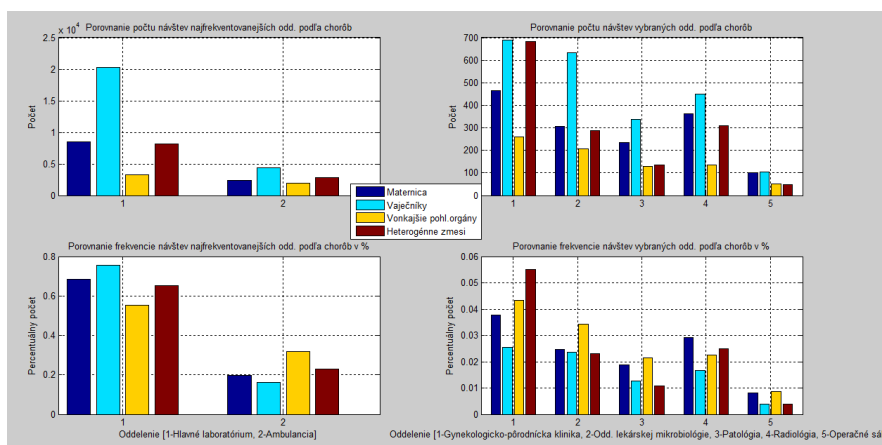
Ide o jednoduchý proces načítania dátového súboru vo formáte .xls, .csv alebo .txt, pričom jednotlivé hodnoty dátového súboru sa uložia do matice.

##### C. Základné štatistiky

Po kliknutí na jedno z tlačidiel základných štatistik sa používateľovi zobrazí grafický výstup s grafmi podľa počtu diagnóz v danom ochorení. Toto grafické okno zahŕňa stĺpcové, bodové alebo koláčové grafy. Na stĺpcových grafoch sa zobrazujú počty prípadov súvisiacich s danou diagnózou na piatich vybraných oddeleniach. Všetky stĺpcové grafy majú na x-ovej osi jednotlivé oddelenia a na osi y-ovej sú jednotlivé početnosti. Koláčové grafy ponúkajú percentuálne zastúpenie jednotlivých diagnóz pri zvolenom type ochorenia. Používateľ si môže zo základného menu vybrať tlačidlo pre štatistiky podľa zvoleného ochorenia a to pre ochorenia maternice, vaječníkov, vonkajších pohlavných orgánov a heterogénnych ochorení.

##### D. Porovnanie návštev oddelení podľa ochorení

Po kliknutí na tlačidlo „Porovnanie návštev oddelení podľa chorôb“ sa používateľovi zobrazí nové grafické okno so štyrmi stĺpcovými grafmi (Obr. 3). Používateľovi sa zobrazí návštevnosť oddelení v reálnych číslach, v hornej časti grafického okna a v percentuálnom vyjadrení v dolnej časti.



Obr. 3 Porovnanie návštevnosti oddelení

##### E. Štatistiky pre počty prechodov medzi oddeleniami

Používateľ pred samotným zobrazením štatistiky musí kliknúť na tlačidlo „Načítanie súboru podľa typu ochorenia“, kde si v m-file zvolí súbor záznamov udalostí. Toto načítanie súboru je

platné pre zobrazenie štatistiky pre počty prechodov medzi oddeleniami, pravdepodobnosti prechodov medzi oddeleniami a pre porovnanie vekovej štruktúry podľa vybraného oddelenia. Po úspešnom načítaní súboru následne používateľ zobrazí potrebné štatistiky po kliknutí na tlačidlo „Štatistiky pre počty prechodov z odd.“. Presné počty prechodov medzi jednotlivými oddeleniami sú premietnuté do grafického výstupu.

#### F. Pravdepodobnosti prechodov medzi oddeleniami

Po kliknutí na tlačidlo „Pravdepodobnosti prechodov medzi odd.“, sa používateľovi zobrazí okno so siedmymi grafmi, ktoré zobrazujú pravdepodobnosti prechodov z vybraného oddelenia na ostatných sedem oddelení. Rovnaký grafický výstup sa zobrazí pre každé ochorenie, ktoré chce používateľ analyzovať. Pre každé oddelenie nám v rámci zvoleného typu ochorenia aplikácia vypočíta s akou pravdepodobnosťou bude prechod z daného oddelenia do nasledujúceho pomocou vzorca:

$$P_{odd_1-odd_2} = \frac{odd_1-odd_2}{sum\_lab}$$

Kde:  $P_{odd_1-odd_2}$  - je pravdepodobnosť prechodu z oddelenia  $odd_1$  do  $odd_2$ ,  
 $odd_1-odd_2$  - je počet prechodov z oddelenia  $odd_1$  do oddelenia  $odd_2$   
 $sum\_lab$  - je súčet počtu všetkých prechodov z oddelenia  $odd_1$  do ostatných oddelení.

#### G. Vplyv veku na vyťaženosť oddelení

Pri gynekologických ochoreniach je pre nemocnice ďalším dôležitým ukazovateľom pri ich analýzach vplyv veku pacientiek. Po kliknutí na tlačidlo „Porovnanie vekovej štruktúry pre zvolené ochorenie“ sa používateľovi zobrazí nové grafické okno s deviatimi stĺpcovými grafmi v percentuálnom vyjadrení.

### V. VYHODNOTENIE A ZÁVER

V rámci navrhovanej aplikácie sme sa zamerali na analýzu ôsmich oddelení. Ešte pred samotnou analýzou zameranou na organizačnú perspektívu sme uskutočnili základné štatistické analýzy, podľa ktorých sme zistili aké percentuálne zastúpenie majú jednotlivé diagnózy v rámci zvolených ochorení, ako aj frekvenciu návštevnosti jednotlivých oddelení pri rôznych ochoreniach. Podľa získaných výsledkov sme identifikovali najviac frekventované oddelenia, ktorými sú *Hlavné Laboratórium* a *Ambulancia*. V ďalšej časti aplikácie sme sa zamerali na analýzu prechodov z jedného oddelenia do druhého pri zvolených oddeleniach. Z pohľadu spolupráce medzi rôznymi oddeleniami sa ako najzaujímavejšie javí oddelenie *Rádiológia*, ktoré má pri všetkých druhoch ochorení značnú mieru spolupráce predovšetkým s oddeleniami *Ambulancie* a *Gynekologicko-pôrodníckej kliniky*. Preto by mohlo byť vhodné navrhnúť kooperáciu medzi týmito oddeleniami aj vzhľadom na to, že určité percento prechodov je smerované z týchto oddelení na oddelenie *Rádiologie*. Podobný trend možno nájsť aj pri oddelení *Patologie*, ktoré má okrem spomínaných dvoch oddelení veľkú mieru spolupráce aj s oddelením *Hlavného Laboratória*. Z pohľadu analýzy jednotlivých ochorení a návštevnosti príslušných oddelení je dôležitý vplyv veku pacientiek, ktorému sme sa venovali v poslednej fáze analýz. Prípady uvedené v záznamoch udalostí sa v prevažnej miere týkajú pacientiek vo veku nad 50 rokov, pričom pri ochoreniach týkajúcich sa vaječníkov a heterogénnych ochoreniach je najpočetnejšou skupinou veková kategória 51-65 rokov, pri ochoreniach týkajúcich sa maternice ešte staršia veková kategória 66-75 rokov.

### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

### ZDROJE

- [1] Process Mining Manifesto. Dostupné na internete: <[http://www.win.tue.nl/ieetfpm/lib/exe/fetch.php?media=shared:process\\_mining\\_manifesto-small.pdf](http://www.win.tue.nl/ieetfpm/lib/exe/fetch.php?media=shared:process_mining_manifesto-small.pdf)>
- [2] Bose, J. CH., Van der Aalst, W. M P., "Analysis of Patient Treatment Procedures." The BPI Challenge Case Study. Eindhoven, 2011.
- [3] Caron, F., Vanthienen, J., De Weerd, J., Baesens, B., "Beyond X-Raying a Care-Flow: Adopting Different Focuses on Care-Flow Mining." The BPI Challenge Case Study. Leuven, 2011
- [4] Varvaressos, G., "Semantic Process Mining. Case Study", BPIC 2011
- [5] 7th International Workshop on Business Process Intelligence 2011. Business Processing Intelligence Challenge (BPIC). august-september 2011. Dostupné na internete: <<http://www.win.tue.nl/bpi2011/doku.php?id=challenge>>

# Podpora vybraných fáz dolovania procesov v Matlabe

<sup>1</sup>Lukáš OŠČIPOVSKÝ, <sup>2</sup>Cecília HAVRILOVÁ, <sup>3</sup>Ján PARALIČ

<sup>1, 2, 3</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>lukas.oscipovsky@student.tuke.sk, <sup>2</sup>cecilia.havrilova@tuke.sk, <sup>3</sup>jan.paralic@tuke.sk

**Abstrakt** — Oblasť dolovania procesov sa v súčasnosti stáva čoraz významnejšou nielen na akademickej pôde, ale aj v komerčnej sfére. Podniky a organizácie si neustále uvedomujú nutnosť analyzovať svoje dáta z rôznych dostupných informačných systémov za účelom získavania nových znalostí a ich využitia v praxi. Cieľom tohto príspevku je uviesť základné existujúce nástroje vhodné pre analýzu procesných dát a tvorbu modelov a tiež popísať návrh a implementáciu vlastného riešenia v MATLAB-e, primárne určeného pre študentov Technickej univerzity v Košiciach. Toto riešenie má slúžiť najmä ako edukatívny nástroj vo výučbe pre pochopenie problematiky dolovania procesov.

**Keľúčové slová** — dolovanie procesov, procesný model, riadenie pracovného toku, záznamy udalostí

## I. ÚVOD

Dolovanie procesov v záznamoch udalostí predstavuje novodobú techniku spracovania a analýzy záznamov udalostí. Táto technológia tvorí pomyselný most medzi dolovaním v dátach a modelovaním biznis procesov [1]. Umožňuje odhaliť procesy zo záznamov udalostí a zobrazit' ich v procesnom modeli tak, aby bolo možné jednoducho overit' ako ľudia alebo procesy skutočne pracujú. Podľa Aalst [2], hlavným cieľom dolovania procesov v záznamoch udalostí je extrahovanie, objavenie, monitorovanie a zlepšovanie procesov vykonávaných prostredníctvom analýzy záznamov udalostí.

Získavanie a uchovávanie záznamov udalostí znamená systematické a chronologické ukladanie udalostí a aktivít vykonávaných prostredníctvom rôznych typov IT systémov [3]. Záznamy udalostí obsahujú informácie ako napr. názov vykonávanej aktivity, začiatok/koniec vykonávania, rola/aktér zodpovedný za vykonávanie danej aktivity a pod. Analýzou takýchto záznamov udalostí a aplikovaním rôznych metód dolovania je možné získať kvalitný model popisujúci priebeh realizovaného procesu. Prostredníctvom získaného modelu je možné identifikovať jednotlivé procesy prebiehajúce v organizácii a pochopiť vzťahy a súvislosti medzi vykonávanými aktivitami [4].

Hlavným cieľom tohto príspevku je uviesť základné softvérové riešenia pre tvorbu procesných modelov z dostupných záznamov udalostí a popísať návrh a implementáciu vlastného riešenia, ako aj jeho overenie vo výučbe.

## II. EXISTUJÚCE RIEŠENIA

*ProM*<sup>1</sup> je modulárne open-source prostredie vyvíjané v jazyku Java. Ide vôbec o prvý nástroj pre analýzu procesov na základe záznamov udalostí. *ProM* umožňuje načítať vstupné záznamy udalostí v špecifických formátoch .xes alebo .mxml. Aktuálne existuje viac ako 120 balíkov s 500 dostupnými rozšíreniami. Na tvorbu procesných modelov využíva heuristické algoritmy, genetické algoritmy, fuzzy prístup a pod. Nástroj umožňuje okrem tvorby procesných modelov aj analýzu záznamov udalostí z pohľadu časovej ako aj organizačnej perspektívy.

*Disco*<sup>2</sup> je komerčný nástroj vyvíjaný spoločnosťou Fluxicon. Nástroj umožňuje import dát vo formátoch .csv, .xes a .mxml. Na prvotnú analýzu dát slúžia základné zobrazenia štatistík pre zvolené atribúty. Procesný model je tvorený prostredníctvom algoritmu fuzzy miner.

<sup>1</sup> <http://www.processmining.org/prom/start>

<sup>2</sup> <http://www.processmining.org/prom/start>



Oba nástroje poskytujú výborné prostredie pre tvorbu procesných modelov a analýzu záznamov udalostí. Ich použitie však vyžaduje pokročilé znalosti tejto problematiky. Z tohto dôvodu sme sa rozhodli navrhnúť a implementovať nástroj, ktorý by bol vhodnejší na výučbové účely, nadväzujúc na znalosti študentov prostredia Matlab, s ktorým pracujú vo viacerých predmetoch a pri záverečných prácach.

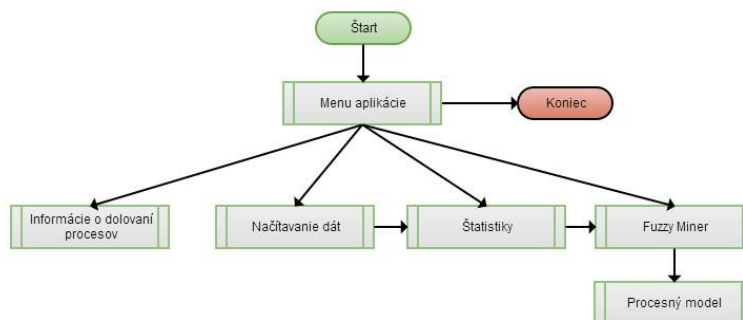
### III. NÁVRH A IMPLEMENTÁCIA RIEŠENIA

#### A. Požiadavky na riešenie

Pre implementáciu nástroja pre dolovanie procesov bolo zvolené vývojové prostredie Matlab<sup>3</sup> od spoločnosti MathWorks. Hlavný dôvod použitia tohto prostredia oproti iným programovacím jazykom spočíval najmä v tom, že ide o prostredie, ktoré je študentom dobre známe z výučby v rámci rôznych predmetov, poznajú ho a vedia s ním pracovať. Navrhované riešenie má byť edukatívne, študentom má vysvetliť základné princípy dolovania procesov a následne im umožniť vyskúšať si experimentovanie v tejto oblasti. Medzi základné požiadavky na funkčnosť aplikácie patrili:

- zobrazenie základných informácií o dolovaní procesov v záznamoch udalostí
- možnosť načítania dátového súboru (.txt, .xls/.xlsx, .csv, .mat) z ľubovoľného miesta na lokálnom disku
- možnosť nastavenia oddeľovača pre načítanie správneho formátu súboru do tabuľky
- výpočet základných štatistík pre jednotlivé prípady záznamov
- výpočet základných štatistík pre jednotlivé aktivity procesu
- výpočet dĺžky trvania aktivít
- výpočet frekvencií kombinácií aktivít a vytvorenie tabuľky následností
- implementácia algoritmu fuzzy miner
- vizualizácia procesného modelu prostredníctvom algoritmu fuzzy miner
- uloženie tabuľky následností do formátu .xls

Hlavnú časť navrhovaného riešenia tvorí samotný algoritmus Fuzzy Miner a jeho zobrazenie pomocou procesného modelu vytvoreného cez anotácie grafov. Celý systém stojí na načítaní dát a ich spracovaní do výsledného modelu. Implementovaný program predstavuje názorne ukážanie dolovania procesov použitím záznamov udalostí a ich transformáciu do procesného modelu. Pre interaktivitu s používateľom bola aplikácia vytvorená pomocou grafického používateľského prostredia (GUI).



Obr. 1 Návrh aplikácie

#### A. Informácie o dolovaní procesov

Po spustení nástroja sa používateľovi zobrazí súhrn informácií o dolovaní procesov. Nachádzajú sa tu základné pojmy ako definícia procesu, záznamov udalostí, procesného modelu, aktivity, procesnej inštancie a pod.

#### B. Načítavanie dátového súboru

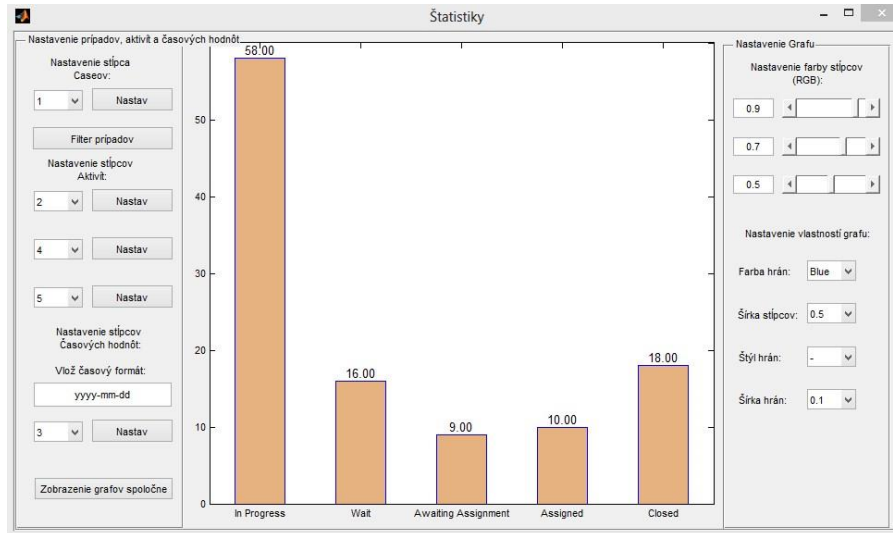
Ďalšou funkciou nástroja je načítavanie dát do tabuliek. Tento proces funguje na princípe algoritmu, ktorý skenuje riadky súboru a pomocou oddeľovača (delimiter) ich vkladá do poľa (matice). Matlab podporuje niekoľko oddeľovačov a tým umožňuje načítať rôzne súbory. Po zadaní týchto parametrov v aplikácii systém načíta súbor do matice a zobrazí ho v podobe tabuľky, ktorú si používateľ môže prezrieť.

<sup>3</sup> <http://www.mathworks.com/products/matlab/>

### C. Štatistiky

Ďalšou dôležitou časťou aplikácie je grafické znázornenie štatistík prípadov, aktivít ako aj časových hodnôt zvolených zo súboru. Po odstránení nepotrebných riadkov pre ďalšiu analýzu (v predchádzajúcom kroku) si používateľ vyberie stĺpce, ktoré chce analyzovať.

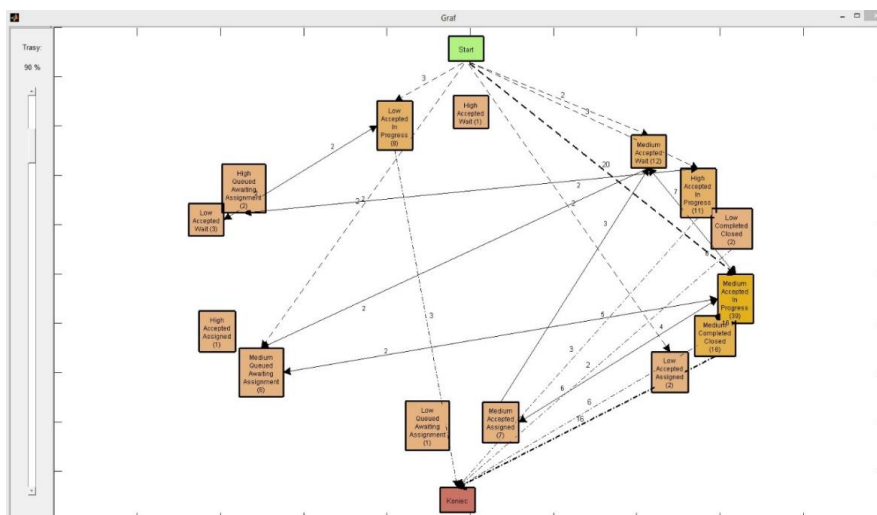
Prvým krokom je zadefinovanie prípadov. Po zadaní čísla stĺpca zo zoznamu sa zobrazí graf relatívnej početnosti prípadov a udalostí. Ak záznam obsahuje časové hodnoty prípadov, je možnosť ich zobrazit' v grafe. Po zadaní stĺpca a formátu dátumu (napr. yyyy-mm-dd) do príslušného textového poľa sa zobrazí stĺpcový graf zobrazujúci minimálnu, priemernú a maximálnu dĺžku prípadov. Všetky štatistiky boli implementované do jedného grafu (Obr. 2).



Obr. 2 Štatistiky

### D. Fuzzy miner

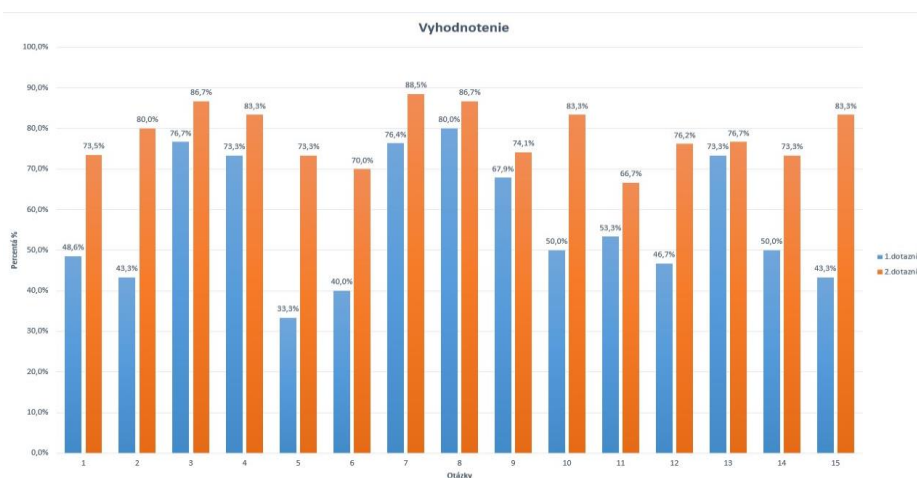
Kľúčovou funkciou aplikácie je algoritmus Fuzzy miner. Po zobrazení štatistík a nastavení stĺpcov pre analýzu je na rade vytvorenie tabuľky následnosti. Tabuľka následnosti je tvorená z frekvencií medzi jednotlivými aktivitami v zázname udalostí. Procesný model z výslednej tabuľky následnosti bol vytvorený cez anotácie grafov. Aplikácia zobrazuje všetky aktivity a ich frekvencie. Z grafu je vidieť ako procesy prebiehajú a ktorá časť procesu je frekventovanejšia ako ostatné alebo opačne. Zobrazenie všetkých trás a hrán v modeli síce predstavuje reálne správanie podľa záznamov ale je neprehľadným riešením. Pre tento dôvod bolo implementované nastavenie zobrazenia frekvencie hrán. Používateľ si môže zobrazit' potrebné množstvo hrán pre analýzu tak, aby videl tie najpodstatnejšie udalosti, ktoré v zázname nastali. Pomocou posuvníka je možné zobrazit' toľko ciest, koľko používateľ potrebuje (Obr. 3). Vzátvorkách anotácií sú zobrazené jednotlivé relatívne početnosti. Čím kontrastnejšia farba aktivity, tým sa vyskytuje v prípadoch častejšie. Hrúbka hrán je tvorená podľa frekvencie z tabuľky následnosti. Štýl hrán je rozdielny pri štartovacích a ukončovacích trasách.



Obr. 3 Procesný model v aplikácii

#### IV. VYHODNOTENIE APLIKÁCIE

Nástroj bol implementovaný pre lepšie pochopenie problematiky dolovania procesov študentmi. Pre túto skutočnosť boli vytvorené dva dotazníky za účelom získania spätnej väzby. Dotazníky boli realizované formou testu pre skupinu študentov tak, aby sa zistilo či pri vysvetľovaní témy pochopili prezentovanú problematiku. Vzorka študentov už mala základné vedomosti z oblasti dolovania procesov aspoň na úrovni jednej prednášky v predmete. Prvý dotazník predstavoval 15 základných otázok o dolovaní procesov. Otázky boli rôzneho typu (dichotomické až po výber z viacerých možností). Dotazník prevažne obsahoval otázky zo základných pojmov (proces, udalosť, aktivita, prípad, atď...) určenej témy tak, aby študenti mohli uspieť aj bez hlbších vedomostí o dolovaní procesov. Druhý dotazník študenti vyplnili po ukázaní a práci s prezentovaným nástrojom. Formulár predstavoval rovnakých 15 testových otázok a ďalších päť predstavovalo spätnú väzbu na nástroj. Na nasledujúcom grafe sú zobrazené výsledky z oboch testov a aj ich vyhodnotenie. Na Obr. 4 je zobrazené percentuálne vyjadrenie otázok z dotazníkov. Pri každej otázke druhého testu došlo k zlepšeniu, pričom odpovede sa zlepšili v rozmedzí od 3% do 40%.



Obr. 4 Vyhodnotenie testov

#### V. ZÁVER

Hlavným cieľom práce bolo vytvoriť edukatívny nástroj, ktorý umožní študentom rýchlo a jednoducho pochopiť problematiku dolovania procesov za použitia zvolených prístupov a algoritmov. Nástroj poskytuje základné štatistiky z dát pre ich lepšie pochopenie a analýzu. V systéme je dostupné zobrazenie relatívnych početností od prípadov až po časové známky tak, aby si používateľ utvoril predstavu o danom zázname. Algoritmus pre vytvorenie procesného modelu je možné použiť pri veľkom množstve aktivít tak, aby zrozumiteľne zobrazil tabuľku následností a následne aj správanie formou procesného modelu. Pre používateľa bolo vytvorené interaktívne prostredie, ktoré zobrazuje možnosti systému ako takého. Aplikácia nástroja do výučby a následné vyhodnotenie zadaných testov potvrdili schopnosť systému podať informácie o dolovaní procesov vo vhodnej a zrozumiteľnej forme.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

#### ZDROJE

- [1] Aalst, Will M.P.: Process mining: Discovery, Conformance and Enhancement of Business Processes. Springer Publishing Company, 2011. ISBN 978-3-642-19344-6.
- [2] Akman, B., Demirors, O.: Applicability of Process Discovery Algorithms for Software Organizations. In: 35th Euromicro Conference on Software Engineering and Advanced Applications. Patras, 2009. ISBN 978-0-7695-3784-9
- [3] Castellanos, M., Casati, F., Dayal, U.: A Probabilistic-based Approach to Process Model Discovery. In: IEEE ICDE Workshops. Hannover, 2011. ISBN 978-1-4244-9196-4
- [4] Fischer, R., Wittern, E., Schneider, J., Tai, S.: Pattern-based Process Discovery. In: 5th IEEE International Conference on Service-Oriented Computing and Applications (SOCA). Taipei, 2012 - ISBN 978-1-4673-4773-0

# Analýza, modelovanie a simulácia skladových procesov

Monika POMIETLOVÁ, Tomáš LOJKA, Iveta ZOLOTOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach

[monika.pomietlova@student.tuke.sk](mailto:monika.pomietlova@student.tuke.sk), [tomas.lojka@tuke.sk](mailto:tomas.lojka@tuke.sk), [iveta.zolotova@tuke.sk](mailto:iveta.zolotova@tuke.sk)

**Abstrakt** — Článok popisuje modelovanie a optimalizáciu skladových procesov. Problematika skladovania a optimalizácie skladov sa úzko dotýka všetkých podnikov, či už väčších alebo menších. Preto sa ňou v poslednom období zaoberá čoraz viac publikácií i odborných prác. U nás i vo svete je to veľmi dôležitá téma zrýchľujúca celý proces výroby a expedície tovarov nájdením úzkych miest a ich následnou optimalizáciou. Dnešný technologický pokrok v oblasti simulovania umožňuje reprezentovať takmer realný model skladu. Na ňom je možné testovať rôzne alternatívy optimalizácie alebo nájdenie numerický vyjdriteľných úzkych miest. Takéto testy v reálne bežiacom sklade by mali za následok spomalenie, prípadne zastavenie skladových procesov.

**Keľúčové slová** — skladovanie, skladové procesy, optimalizácia, Arena softvér

## I. ÚVOD

Témou tohto článku je analýza, modelovanie a simulácia skladových procesov. Prevádzkovatelia skladov dnes riešia omnoho viac ako kedykoľvek predtým [1]. Potrebujú efektívny nástroj na riadenie skladových procesov a vyhodnotenie svojej práce [2]. Takýmto nástrojom sa môže stať simulačný softvér. V mojom prípade som si vybrala softvér Arena, pretože je jednoducho použiteľný pre rôzne typy skladov aj bez potreby ovládania akéhokoľvek programovacieho jazyka na rozdiel od ostatných dostupných softvérov. Preto si myslím, že je vhodný na jednoduché, no pritom účinné modelovanie a simuláciu všetkých procesov prebiehajúcich v sklade, na ich pochopenie a nájdenie vhodných riešení všetkých problémov vyskytujúcich sa počas tohto procesu [3].

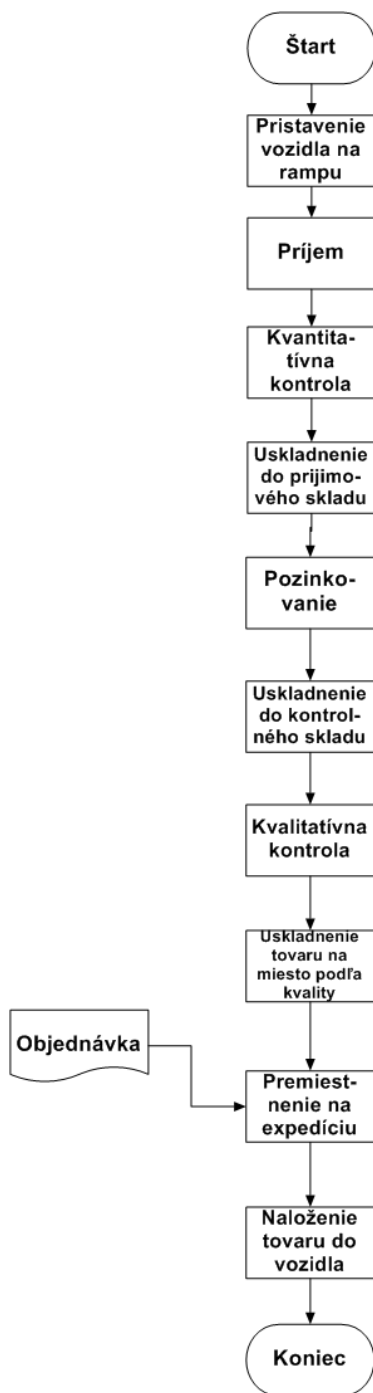
## II. KEĽÚČOVÉ VÝCHODISKA PRE SKLADOVANIE A JEJ OPTIMALIZÁCIU

Hlavným východiskom pre analýzu skladovania a jej optimalizáciu je poznať typ skladu a kam je ho potrebné v delení skladov zaradiť [4]. Možeme ich rozdeliť podľa mechanizácie, stavebného prevedenia, výšky objektu, podľa prietoku tovaru [4],[5]. Okrem toho podnik musí vyhodnotiť všetky známe faktory ovplyvňujúce usporiadanie skladu a vytvoriť tak pre seba optimálne skladové procesy, pri ktorých sa využije skladový priestor čo najefektívnejšie [6].

## III. NÁVRH MODELU SKLADU

Návrh modelu skladu je proces, ktorý predstavuje tok materiálu. Je to prehľadný spôsob ako navrhnuť proces pred jeho samotným simulovaním [7]. Tento model bude slúžiť ako podklad pre modelovacie prostredie Arena. Skladá sa z niekoľkých krokov:

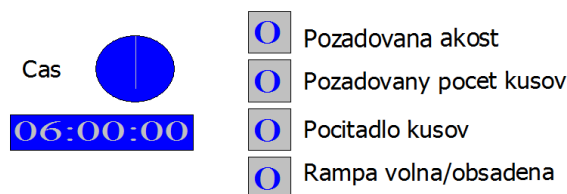
- Z valcovne privezie nepozinkované zvitky vlak na príjmovú rampu, kde sa vyložia.
- Privezené zvitky sa príjmu na sklad, prepočítajú a uskladnia sa v príjmovom sklade.
- Z príjmového skladu sa presunú na pozinkovacia linku, kde sa pozinkujú a uskladnia sa do dočasného kontrolného skladu.
- V kontrolnom sklade sa vyhodnotí kvalita pozinkovaných zvitkov, rozdelia sa na 3 akosti a uložia na miesto pre nich vyhradené.
- Ak príde objednávka, požadovaná akosť a množstvo sa presunú na expedíciu.
- Skompletizovaná objednávka sa z vyskladňovacej rampy vyloží do kamióna pripraveného na expedíciu [1].



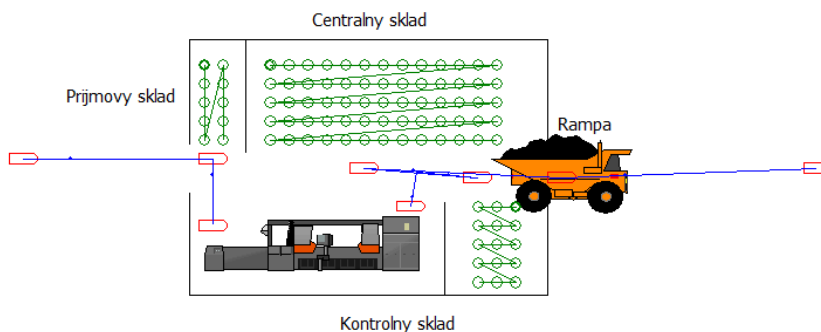
Obrázok 1: Návrh modelu skladu

#### IV. SKLAD V MODELOVACOM PROSTREDÍ ARENA

Pred samotnou optimalizáciou je vytvorený prvotný sklad. Ten slúžil ako východiskový a porovnávací bod v procese optimalizácie [8]. Sklad, pozostávajúci z 90 uskladňovacích miest, je rozdelený na 3 časti: príjmovú (10 miest), kontrolnú (15 miest) a centrálnu (65 miest). Vzhľadom na obmedzený počet modulov a entít v študentskej verzii programu Arena do procesu počas jeho priebehu vstupuje 100 kusov zvitkov a 40 objednávok. Počet pracovníkov je tiež obmedzený, preto sa v procese nachádzajú len 4 typy pracovníkov: Baliaci pracovník, kontrolór kvality, manipulačný pracovník vysokozdvížného vozíka a manipulačný pracovník žeriavu. Pozinkované zvitky sa uskladňujú do spoločného centrálného skladu odkiaľ sú následne podľa akosti vyskladnené na expedíciu.



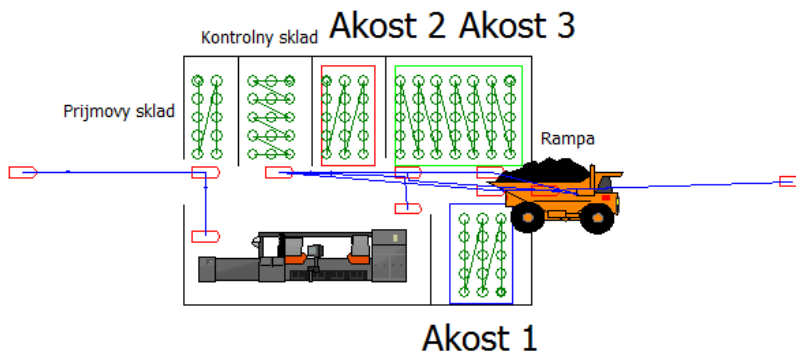
**Obrázok 2: Panel aktuálnych informácií**



**Obrázok 3: Vizualizácia - Prvotný sklad**

*A. Optimalizácia pomocou ABC metódy a pevného uskladnenia*

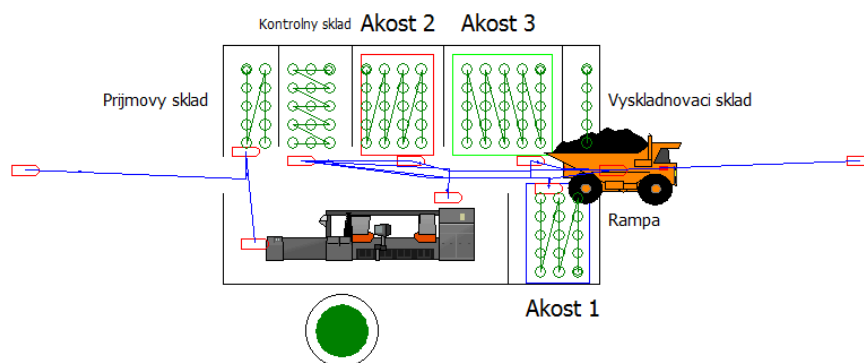
V tomto sklade (ďalej len Sklad\_1) je použité blokové skladovanie zvitkov. V tomto prípade sú uložené podľa metódy pevného ukladania, pretože každá akost' má svoje stále miesto v sklade. Akost' 1 má najvyššiu prioritu, je najbližšie pri pozinkovacej linke, ale aj čím bližšie ku vyskladňovacej rampe, aby sa zvitky tejto najlepšej akosti deformovali manipuláciou čo najmenej. Zároveň sú uskladnené podľa metódy ABC, pretože akost' 3 má najvyššiu frekvenciu objednávok, tak je uložená čo najbližšie ku vyskladňovacej rampe. Sklad je rozdelený na 5 rôznych častí: príjmový sklad (10 miest), kontrolný sklad (15 miest), sklad akosti 1 (15 miest), sklad akosti 2 (15 miest) a sklad akosti 3 (35 miest).



**Obrázok 4: Vizualizácia – Sklad\_1**

*B. Optimalizácia pomocou ABC metódy, pevného a prípravného uskladnenia*

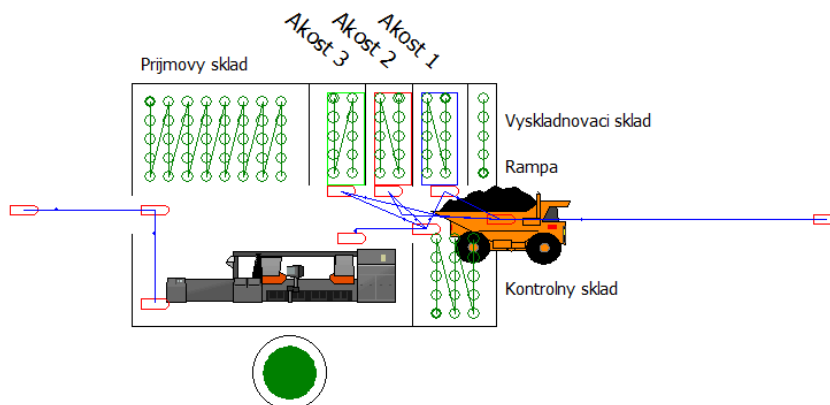
V tomto sklade (ďalej len Sklad\_2) je použité blokové skladovanie zvitkov, uskladnenie podľa metódy pevného ukladania, tak ako v predchádzajúcom prípade. Novou optimalizačnou metódou oproti Sklad\_1 je prípravné uskladnenie, ktoré sa prejavuje vznikom vyskladňovacieho skladu. Tento síce zaberá časť kapacity skladu akosti 3, ale zrýchli celý priebeh procesu. Sklad je rozdelený na 6 rôznych častí: príjmový sklad (10 miest), kontrolný sklad (15 miest), sklad akosti 1 (15 miest), sklad akosti 2 (15 miest), sklad akosti 3 (30 miest) a vyskladňovací sklad (5 miest).



Obrázok 5: Vizualizácia – Sklad\_2

### C. Optimalizácia pomocou metódy Just-in-time

V tomto sklade (ďalej len Sklad\_3) je použité blokové skladovanie zvitkov, uskladnenie podľa metódy pevného a prípravného uskladnenia, tak ako v predchádzajúcom prípade. Novou optimalizačnou metódou oproti Sklad\_2 je Just-in-time metóda, ktorá je aplikovaná zväčšením kapacity príjmového skladu na úkor priestorov pre jednotlivé akosti. Táto optimalizácia spočíva v naskladnení veľkého počtu nepozinkovaných plechov do príjmového skladu, odkiaľ sa na ďalšie spracovanie vyberú až po príchode objednávky na jednotlivé akosti, čo zrýchli celý proces skladovania oproti predošlému modelu. Taktiež vyrieši problém ukladania veľkého množstva rôznych akostí na rôzne miesta v sklade, pretože väčšina kapacity skladu je zaplnená ešte nepozinkovanými zvitkami. Sklad je rozdelený na 6 rôznych častí: príjmový sklad (40 miest), kontrolný sklad (15 miest), sklad akosti 1 (10 miest), sklad akosti 2 (10 miest), sklad akosti 3 (10 miest) a vyskladňovací sklad (5 miest).



Obrázok 6: Vizualizácia – Sklad\_3

### D. Porovnanie optimalizačných metód

V prvotnom sklade dosahoval priemerný čas vybavenia objednávky 4,4100 hodiny, čo je pre reálne fungovanie skladu neprípustné. Použitím metódy pevného ukladania (vyhradenie miesta pre jednotlivé akosti) a metódy ABC (umiestnenie najfrekventovanejšie objednávané akosti ku vyskladňovacej rampe) sa dosiahlo zlepšenie priemerného času vybavenia objednávky na 3,6828 hod., t. j. zlepšenie o 0,7272 hod., teda o 43,632 minút. Toto zlepšenie znamená skrátenie času potrebného na vybavenie objednávky o 16,49 %.

Pridaním metódy prípravného uskladnenia sa dosiahlo zlepšenie priemerného času vybavenia objednávky na 2,9132 hod., t. j. zlepšenie o ďalších 0,7696 hod., teda o ďalších 46,176 minút. Toto zlepšenie znamená skrátenie času potrebného na vybavenie objednávky o 20,90 %. Oproti prvotnému skladu nastalo zlepšenie o 33,94 %.

Nakoniec pridaním metódy Just-in-time sa dosiahlo zlepšenie priemerného času vybavenia objednávky na 1,7515 hod., t. j. zlepšenie o ďalších 1,1617 hod., teda o ďalších 69,702 minút. Toto zlepšenie znamená skrátenie času potrebného na vybavenie objednávky o 39,88 %. Oproti prvotnému skladu nastalo zlepšenie o 60,28 %.

Obe posledné metódy dosahujú časy vybavenia objednávok prijateľné v reálnom sklade. Metóda prípravného uskladnenia je síce pomalšia, ale na druhej strane bezpečnejšia ako Just-in-time. Vytvára väčšie zásoby jednotlivých akostí zvitkov, ktoré sú vždy pripravené na



vyskladnenie. Na rozdiel od tejto metódy, metóda Just-in-time je rýchlejšia, ale vzniká riziko, že pri príchode objednávky nebude k dispozícii dostatočný počet požadovaných zvitkov. Je len na podniku, pre ktorú metódu sa rozhodne. Prípravné uskladnenie je bezpečné, pohodlné a relatívne rýchle. Metóda Just-in-time je vhodná len pre podniky kde sú dodávateľské reťazce na dostatočnej úrovni. V nasledujúcej tabuľke sú znázornené výsledky získane simuláciou 4 variant skladov a 3 objednávok., z čoho vychádza aj uvedený záver v tomto odstavci.

	Prvotný sklad	Sklad_1	Sklad_2	Sklad_3
<b>Objednávka 1</b>	5,2998	2,1498	1,6632	1,6515
<b>Objednávka 2</b>	3,8054	4,5365	3,0132	1,9081
<b>Objednávka 3</b>	4,1248	4,3621	4,0632	1,6948
<b>Priemer</b>	4,4100	3,6828	2,9132	1,7515

**Tabuľka 1: Porovnanie časov vybavenia objednávok v hodinách**

## V. ZÁVER

Problematika optimalizácie skladov sa stáva z roka na rok dôležitejšou a rozoberanejšou otázkou [9]. V dnešnej hektickej dobe je veľmi dôležité uspokojiť svojich odberateľov načas, s prehľadom a hlavne efektívne. Dnes si žiadny podnik nemôže dovoliť čakať na dodanie materiálu alebo tovaru. Celý proces výroby a distribúcie musí byť plynulý. Práve preto musia byť každým skladovom procese moderného podniku odhalené včas všetky úzke miesta a odstránené pomocou optimalizačných metód [10]. Pri aktuálnom technologickom pokroku už nie je potrebné skúšať rôzne alternatívy v reálnom sklade, čo má za následok spomalenie alebo aj neúspešné pokusy. Je oveľa jednoduchšie a efektívnejšie namodelovať celý skladový proces pomocou dostupných simulačných softvérov, vďaka ktorým sa stane prehľadným a všetky problémy jednoduchšie riešiteľnými. Toto preukázalo aj moje modelovanie a optimalizácia skladu pozinkovacej haly v U. S. Steel Košice, kde sa mi použitím rôznych optimalizačných metód v simulačnom prostredí podarilo znížiť čas potrebný na vybavenie objednávky o viac ako 60 %.

## POĎAKOVANIE

Táto práca bola podporená grantom z KEGA - 021TUKE-4/2012 (70%) a grantom VEGA - 1/0286/11 (30%).

## POUŽITÁ LITERATÚRA

- [1] E. Spišák, J. Hudák, J. Tkáčová, J. Tkáč, "Pozinkované plechy v automobilovom priemysle a ich hodnotenie". Košice : Technická univerzita v Košiciach, Strojnícka fakulta, Katedra technológií a materiálov.
- [2] J. Hašek, "Logistika skladov, skladovania a riadenia zásob". Pardubice, Univerzita Pardubice – Dopravná fakulta Jana Pernera. 2009. 62. s.
- [3] K. Viestová, a kolektív : "LEXIKÓN LOGISTIKY". Vyd. Bratislava: Iura Edition, spol. s.r.o., 2007, s.163, ISBN 978-80-8078-160-6
- [4] K. Halabíková, "Skladové hospodárstvo konkrétneho podniku". Brno : Masarykova univerzita. 2010, . s. 59
- [5] K.B. Ackerman, "Practical Handbook of Warehousing – The Functions of Warehouse". 1990, s. 19-30
- [6] J. Šaderová, "Tok tovaru logistickým reťazcom veľkoobchodu. Košice". TU Košice, Fakulta BERG. ISSN 1451-107X, s. 33-37
- [7] D.M. Lambert, J.R. Stock, L.M. Ellram, "Logistika." 1. vyd., Praha : Computer Press, 2000. ISBN 80-7226-221-1
- [8] J. Sixta, V., Macat, "Logistika – teória a prax". 1.vyd. Brno: Computer Press, 2005. ISBN 80-251-0573-3
- [9] Ch. Schulte, "Logistika". 1. vyd. Praha: Victoria Publishing, 1994. ISBN 80-85605-87-2
- [10] A. Stehlik, "Logistika – strategický faktor manažérskeho úspechu". 1. vyd. Brno: Studio Kontrast, 2002. ISBN 80-238-8332-1



# Prepínače v Hybridných RF/FSO Komunikačných Systémoch

<sup>1</sup>Peter STRNISKO, <sup>2</sup>Luboš OVSENÍK, <sup>3</sup>Ján TURÁN

<sup>1,2,3</sup> Katedra Elektroniky a Multimediálnych Telekomunikácií, Fakulta Elektrotechniky  
a Informatiky, Technická Univerzita Košice, Slovenská Republika

<sup>1</sup>ing.peter.strnisko@gmail.com, <sup>2</sup>lubos.ovsenik@tuke.sk, <sup>3</sup>jan.turan@tuke.sk

**Abstrakt** — V tomto článku sú opísané vlastnosti hybridného RF/FSO komunikačného systému. Takisto sú tu uvedené vlastnosti primárnej a záložnej RF linky a rozdiely medzi prepínaním počas komunikácie medzi jednotlivými linkami. Časť článku je venovaná Raptorovým kódom krátkej dĺžky. Tieto Raptorové kódy sú vyvinuté na dosiahnutie vysoko účinného soft switchingu v hybridných komunikačných systémoch.

**Kľúčové slová** — Hard switching, hybridný RF/FSO komunikačný systém, soft switching, záložná RF linka.

## I. ÚVOD

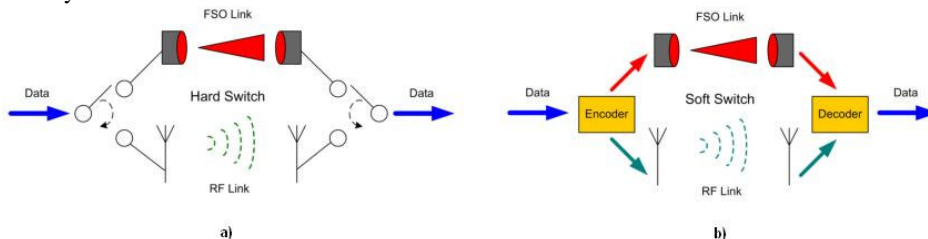
Hybridný komunikačný systém je zložený z dvoch komunikačných systémov, kde primárna komunikácia medzi vysielačom a prijímačom prebieha cez FSO systém. V prípade zlyhania tejto linky sa využíva druhá, takzvaná záložná RF linka. Neoddeliteľnou súčasťou celého komunikačného systému je prepínač, ktorý slúži na prepínanie medzi primárnou a záložnou linkou. Slúži taktiež na spätné prepnutie liniek, ak sa obnoví spojenie na primárnej linke.

Tento článok je zameraný na opis základných vlastností primárnej a záložnej linky. Takisto sú tu popísané typy používaných prepínačov medzi týmito dvoma linkami a ich výhody a nevýhody.

## II. PREPÍNAČE PRE HYBRIDNÉ RF/FSO KOMUNIKAČNÉ SYSTÉMY

Optická komunikácia voľným prostredím – FSO (Free Space Optics) dosahuje vysokú rýchlosť prenosu dát pri nízkej zložitosti systému, ale na druhej strane je silne ovplyvnená atmosférickými stratami kvôli hmle a iskreniu. Rádio frekvenčný – RF (Radio Frequency) systém má nižšiu rýchlosť prenosu dát ako FSO systém, ale je menej citlivý na počasie. Hybridný RF/FSO komunikačný systém spája výhody oboch systémov a tým vytvára komunikačný systém, ktorý je relatívne necitlivý na zmeny počasia. Prepínanie medzi dvomi linkami hybridného systému zabezpečuje prepínač. V súčasnej dobe môžeme použiť dva typy prepnutí medzi použitou primárnou a záložnou linkou a to buď „Hard switching“, alebo „Soft switching“. Technika hard switchingu je dosť neefektívna, z toho dôvodu, že stále iba jedna z liniek je aktívna. V prípade soft switchingu sú aktívne súčasne obidve linky. Z doteraz dostupnej literatúry je jasné, že existuje technika soft switchingu, založená na použití Raptorových kódov krátkej dĺžky. Pakety kódované Raptorovými kódmi sú vysielané na obidvoch linkách, primárnej aj záložnej súčasne [1]. Na Obr. 1 sú dve možné konfigurácie hybridného RF/FSO komunikačného systému. Prvá časť Obr.1, časť a) zobrazuje konfiguráciu hard switchingu, kde medzi vysielačom a prijímačom je vytvorené spojenie iba cez jeden z dvoch kanálov a to buď cez FSO kanál, alebo cez RF kanál. Systém s vysokou rýchlosťou prenosu dát, čiže FSO systém je zvolený len vtedy, ak podmienky kanála umožňujú spoľahlivú komunikáciu vzhľadom na poveternostné podmienky. V opačnom prípade všetky dáta sú odosielané cez RF kanál. Nesmieme zabudnúť na to, že vysielač a prijímač musia byť koordinované prostredníctvom spätnej väzby na výber toho správneho kanála na prenos informácií. Hlavnou nevýhodou tejto konfigurácie je, že vždy jedna z dvoch liniek je nečinná. V praxi to znamená, že väčšinou je nečinná na prenos dát rýchlejšia FSO linka, kvôli iskreniu a dáta sa prenášajú cez pomalšiu, čiže záložnú RF linku [3]. Z toho dôvodu vznikla myšlienka soft switchingu. Táto konfigurácia je znázornená v druhej časti Obr.1, v časti b) [4]. Pri soft

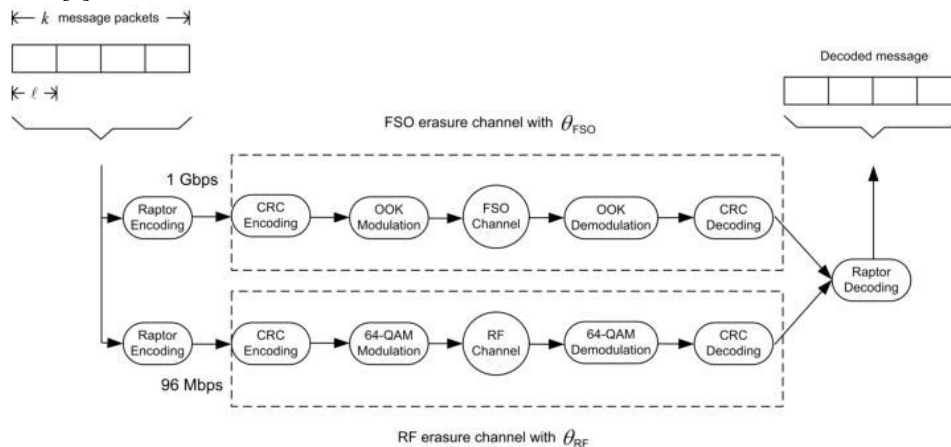
switchingu dáta sú kódované jednoduchým LDPC (Low Density Parity Check) kódom, kde sa kódové slovo rozdelí a časť kódového slova prechádza cez FSO systém a časť cez RF systém a rýchlosť prenosu dát závisí od okamžitých podmienok kanála. Táto technika prepínania je lepšia v porovnaní s predchádzajúcou, ale podmienky prenosového kanála pri prenose dát musia byť známe aj pri tejto konfigurácii ako pre vysielateľ, tak aj pre prijímač. Preto sa hľadali kódy, pri použití ktorých sa nevyžaduje znalosť podmienok kanála vo vysielateľi a sú schopné sa prispôsobiť rýchlosťou prenosu dát, podmienkam kanála. Tejto podmienke vyhovujú Raptorové kódy [1], ktoré sa prispôbujú rýchlosťou prenosu dát podmienkam kanála, len spätnou väzbou jedného bitu na správu. Ako už bolo spomenuté vyššie, pri použití Raptorových kódov kóder nevyžaduje informáciu ohľadom podmienok prenosového kanála a kód sa automaticky prispôbi rýchlosti FSO alebo RF systému v závislosti od poveternostných podmienok. Pri použití LDPC kódov bola potrebná rozsiahla spätná väzba z vysielateľa do prijímača na prispôbenie a synchronizáciu kódov na koncoch obidvoch liniek. Simulácie a výsledky na reálnych modeloch sú bližšie opísané a uvedené v [1], kde sa robili simulácie v troch Kanadských mestách.



Obr. 1 Konfigurácia hybridného RF/FSO systému a) Hard switching b) Soft switching [1]

#### A. Raptorové kódy krátkej dĺžky

Kódovacia technika na úrovni paketov pomocou Raptorových kódov je vyvinutá na dosiahnutie vysoko účinného soft switchingu v hybridných RF/FSO systémoch. Postupnosti kódovaných výstupov sú generované z informačných paketov (Obr. 2) a odosielané súčasne cez linky FSO aj RF. Neexistuje žiadna koordinácia medzi týmito dvomi linkami a pakety sú odoslané buď cez FSO alebo RF kanál s ich príslušnými prenosovými rýchlosťami. Prijímač pozbiera pakety a celá prenesená informácia je dekodovaná, keď je prijatý dostatočný počet paketov [1].



Obr. 2 Soft switching hybridný RF/FSO systém použitím Raptorových kódov [1]

Raptorové kódy sú náhodné dvojdielne grafové kódy, v ktorých každý kódovaný paket je lineárnou kombináciou náhodného množstva informačných paketov. Informácia sa skladá z  $k$  paketov o dĺžke bitu. Definujme  $m$ , ako počet prijatých paketov, potrebných pre úspešné dekodovanie správy. Keďže kóder Raptorových kódov generuje pakety náhodným spôsobom, aj  $m$  je náhodný. Definujme predpokladané pretečenie Raptorového kódu,  $\mathcal{E}_m$  ako:

$$\langle m \rangle = (1 + \mathcal{E}_m)k \quad (1)$$

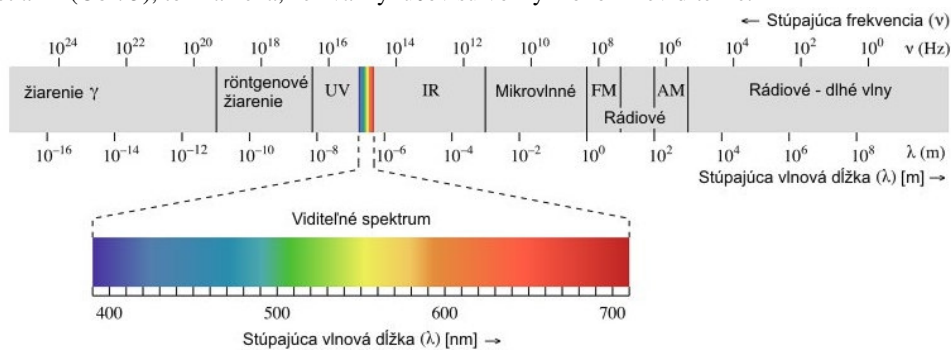
kde  $\langle \cdot \rangle$  označuje predpoklad. Štatistické pretečenie,  $\mathcal{E}_m$  je vlastné Raptorovmu kódu a nezávisí od podmienok prenosového kanála [1].

### III. HYBRIDNÝ RF/FSO KANÁLOVÝ MODEL

Kanáľ je dôležitou časťou každého prenosového systému. Pri návrhu kanálového modelu hybridného RF/FSO komunikačného systému je treba brať do úvahy zvlášť vlastnosti FSO komunikačného systému a zvlášť vlastnosti RF systému. Rozdiely v jednotlivých modeloch sú hlavne v prenosovej rýchlosti jednotlivých typov kanálov a v tom ako atmosféra ovplyvní tieto dva typy komunikačných systémov. Kým FSO komunikačný systém je najviac tlmený v dôsledku hmly, RF komunikačný systém je citlivý na dažď a mokrý sneh. Treba brať do úvahy aj to, že obidva typy kanálov sú tlmené iskrením. Návrhu, parametrom a opisu vlastností jednotlivých typov kanálových modelov sa tento článok podrobne nezaobera, keďže všetky tieto uvedené skutočnosti sú dosť podrobne opísané v [5]. Uvedená [5] sa podrobne venuje hybridnej signalizácii a jednotlivým typom útlmu a výpočtu útlmových koeficientov pre rôzne poveternostné podmienky, podľa známych modelov, akými sú: Kruse, Kim a Al-Naboulsi modely.

### IV. ZÁKLADNÉ VLASTNOSTI PRIMÁRNEJ LINKY

Linka FSO v hybridnom RF/FSO komunikačnom systéme je označovaná za primárnu linku. Je to širokopásmová komunikačná technológia, kde musí byť dodržaná priama viditeľnosť – LOS (Line – of – Sight). Pri tomto komunikačnom systéme sa svetelné impulzy prenášajú úzkym zväzkom lúča cez atmosféru. FSO systémy pracujú v infračervenej oblasti svetla (Obr. 3), to znamená, že zväzky lúčov sú voľným okom neviditeľné.



Obr. 3 Frekvenčné spektrum [6]

Využitie FSO systémov je rôzne. Môže sa použiť na miestach, kde končí vedenie optických vlákien a signál je nutné prenášať ďalej k užívateľom do vzdialenosti približne od 50 m do niekoľko kilometrov. FSO systém je taktiež možné využiť ako záložnú linku pre optickú sieť, aby v prípade nedostupnosti optickej siete nedošlo k strate dát. Ďalej sa FSO systém môže použiť ako dočasná linka, alebo všade tam, kde je komplikovaný terén a nie je možnosť využiť vláknovú optiku. Najväčšou nevýhodou FSO systémov je to, že sú silne ovplyvnené atmosférickými vplyvmi a to najmä hmlou. Už len mierna hmla dokáže spôsobiť značný útlm prenášaného signálu. Ďalším faktorom, ktorý spôsobuje útlm je iskrenie. Na druhej strane vplyv dažďa a snehu na FSO systém je zanedbateľný. Podrobný opis toho, ako atmosféra ovplyvní FSO systém aj s vyjadrením útlmových koeficientov je uvedený v [5].

### V. ZÁKLADNÉ VLASTNOSTI ZÁLOŽNEJ LINKY

Záložná linka pre hybridný RF/FSO systém je dôležitá pre prípad nedostupnosti primárnej linky. Ako už bolo spomenuté vyššie dostupnosť primárnej linky je silne ovplyvnená niektorými atmosférickými vplyvmi a bez záložnej linky by mohlo dôjsť k strate prenášanej informácie. Na tento účel slúži RF linka, ktorá je rádiová linka a nie je ovplyvnená hmlou. Za najvhodnejšie záložné linky z hľadiska dostupnosti sú považované linky pracujúce pod frekvenciou 10 GHz, kde vplyv dažďa na prenášaný signál je relatívne malý. Avšak toto riešenie nie je ľahko dosiahnuteľné. Problém je v tom, že rýchlosť prenášanej informácie na záložnej linke by mala byť aspoň 100 Mbit/s alebo aj vyššia. Bolo by dobré, keby frekvenčné pásmo záložnej linky bolo bezlicenčné, keďže FSO systémy pracujú v bezlicenčnom pásme [2]. Z dostupnej literatúry je známe, že pre záložné linky sa používajú napr. tieto frekvencie: 2,4 GHz, 5 GHz (nevýhodou týchto frekvencií je veľké preťaženie pásma), 10 GHz, 50 GHz a 60 GHz. V nasledujúcej časti budú opísané základné vlastnosti 10 GHz - ovej záložnej linky a 60 GHz – ovej záložnej linky.

#### A. Základné vlastnosti 10 GHz RF linky

Záložná linka musí byť imúnna voči poveternostným vplyvom, aby v prípade nedostupnosti

primárnej linky bola k dispozícii na prenos informácií. Tieto linky pracujú v oblasti milimetrových vln, kde prenos dát je ovplyvnený dažďom, pretože vlnová dĺžka nosnej frekvencie je porovnateľná s veľkosťou dažďovej kvapky. Veľkosť dažďových kvapiek je od 0,1 mm do 10 mm a preto dochádza k rozptylu milimetrových vln, hlavne pre nosné frekvencie okolo 10 GHz. Hmla tieto linky neovplyvňuje, pretože veľkosť častočiek hmly je približne od 1  $\mu$ m do 20  $\mu$ m (preto dochádza pri hmle k útlmu na FSO linke). K útlmu obidvoch liniek hybridného vedenia súčasne nedochádza, pretože keď prší, dažďové kvapky absorbujú hmlové kvapôčky. Výhodou záložných liniek, ktoré pracujú v oblasti milimetrových vln je to, že dokážu spĺňať funkciu záložnej linky od relatívne krátkych vzdialeností až do niekoľko kilometrov [7].

#### B. Základné vlastnosti 60 GHz RF linky

Od roku 2001 je zavedené 7 GHz pásmo, ktoré je bezlicenčné. Toto pásmo sa nachádza medzi frekvenciami 57 GHz a 64 GHz. Toto pásmo sa nachádza v pásme milimetrových vln. Táto časť spektra bola do veľkej miery nevyužitá pre komerčné bezdrôtové aplikácie. V tomto pásme je možné dosiahnuť vysokú prenosovú rýchlosť, je odolne voči rušeniu a takisto je tu možnosť opätovného využívania frekvencie. 60 GHz frekvenčné pásmo bolo využívané armádou pre satelitné komunikácie pre svoju vysokú bezpečnosť. Veľká výhoda tohto frekvenčného pásma súvisí s absorpciou elektromagnetickej energie molekulami kyslíka. Táto absorpcia tlmí signály na diaľku, z toho vyplýva, že prenos signálu na dlhú vzdialenosť za prijímač je nemožný. Toto frekvenčné pásmo sa používa väčšinou na prenos dát do vzdialenosti približne 1 km. Výhodou tejto absorpcie je aj to, že sa žiarenia dvoch 60 GHz liniek v rovnakej oblasti neovplyvňujú, z toho vyplýva aj fakt opätovného využívania frekvencie. Tieto linky môžu pracovať s linkami s bezlicenčným frekvenčným pásmom, ktoré nie sú ovplyvnené absorpciou kyslíka. Pri frekvencii 60 GHz je šírka zväzku lúča oveľa užšia ako pre nižšie frekvenčné pásma. Kým pri frekvencii 2,4 GHz je šírka zväzku lúča 117 stupňov, pri frekvencii 60 GHz je šírka zväzku lúča len 4,7 stupňa [2].

## VI. ZÁVER

Tento článok je venovaný hybridnému RF/FSO komunikačnému systému, kde sú popísané základné vlastnosti primárnej a záložnej linky. Takisto sú tu opísané typy prepínačov medzi týmito dvomi linkami. Ako už vieme, hard switching má jednu veľkú nevýhodu oproti soft switchingu a tou nevýhodou je, že stále jedna z dvoch liniek je neaktívna. Preto viac sa tento článok zaoberá soft switchingom a to pomocou Raptorových kódov. Používajú sa Raptorové kódy krátkej dĺžky ( $k = 16 - 1024$ ). Na základe dostupných výsledkov simulácií, ktoré už boli realizované vychádza, že použitie Raptorových kódov je praktické a účinne v simulovaných prenosových kanáloch. Posledná časť článku je venovaná opisu vlastností 10 GHz a 60 GHz záložnej linky, ktoré sa dosť často používajú v hybridných RF/FSO komunikačných systémoch.

## POĎAKOVANIE

Táto práca bola podporovaná Kultúrnou a edukačnou grantovou agentúrou (KEGA) Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky na základe projektu č. „063TUKE-4/2013 - Použitie diaľkovo ovládaného optického vláknového refraktometra vo výučbe“ a projektu č. „006TUKE-4/2014 - Experimentálny model PON siete TUKE a jeho použitie vo výučbe“. Táto práca je tiež výsledkom zmluvy „APVV-0025-12 - Predchádzanie vplyvu stochastických mechanizmov vo vysokorýchlostných plne optických sieťach“ podporovanej Agentúrou na podporu výskumu a vývoja (APVV) Slovenskej republiky.

## REFERENCIE

- [1] W. Zhang, S. Hranilovic, C. Shi, "Soft - Switching Hybrid FSO/RF Link Using Short - Length Raptor Codes: Design and Implementation" IEEE Journal on selected areas in Communications, Vol. 27, No. 9, December 2009
- [2] P. Mišenčík, L. Ovseník, J. Turán, "Vlastnosti Hybridnej FSO/RF Linky so 60 GHz RF Záložnou Linkou", EEI III. Proceeding of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, 2012
- [3] A. Akbulut, H. G. Ilk, F. Ari, „Design, Availability and Reliability Analysis on an Experimental Outdoor FSO/RF Communication System“, Proc. IEEE Int. Conf. Transparent Opt. Networks, pp. 403 - 406, 2005
- [4] S. Vangala, H. Pishro-Nik, "Optimal Hybrid RF-Wireless Optical Communication for Maximum Efficiency and Reliability," Proc. IEEE Annu. Conf. Inform. Sci. Syst., pp. 684-689, 2007
- [5] P. Strnisko, "Hybridné RF/FSO Komunikačné Systémy (Unpublished work style)", unpublished
- [6] [http://sk.wikipedia.org/wiki/Elektromagnetick%C3%A9\\_%C5%BEiarenie](http://sk.wikipedia.org/wiki/Elektromagnetick%C3%A9_%C5%BEiarenie), 26.06.2014
- [7] G. Singh, T. Singh, Vinaykant, V. Kaushal, „Free Space Optics: Atmospheric Effects & Back Up“, International Journal of Research in Computer Science sISSN 2249 - 8265, Volume 1, Issue 1, (2011), pp. 25 - 30

# Hybridné RF/FSO Komunikačné Systémy

<sup>1</sup>Peter STRNISKO, <sup>2</sup>Luboš OVSENÍK, <sup>3</sup>Ján TURÁN

<sup>1,2,3</sup> Katedra Elektroniky a Multimediálnych Telekomunikácií, Fakulta Elektrotechniky a Informatiky, Technická Univerzita Košice, Slovenská Republika

<sup>1</sup>ing.peter.strnisko@gmail.com, <sup>2</sup>lubos.ovsenik@tuke.sk, <sup>3</sup>jan.turan@tuke.sk

**Abstrakt** — Optické komunikácie voľným prostredím sú uvedené v tomto článku. Okrem všeobecného úvodu sa tu popisuje kanálový model a hybridná signalizácia. Ďalšia časť je venovaná vplyvu počasia na hybridné systémy. Je tu uvedený útlm spôsobený hmlou všeobecne a útlm spôsobený hmlou na GHz vedenia. Posledná časť popisuje vplyv snehu a dažďa na hybridné systémy.

**Príušové slová** — Hybridný RF/FSO komunikačný system, hybridná signalizácia, kanálový model, útlm.

## I. ÚVOD

Optická bezdrôtová komunikácia je rozvíjajúca sa oblasť a priniesla veľké množstvo riešení zložitých komunikačných problémov. Patrí k nim vysoká rýchlosť prenosu dát, vysoký výkon a minimálne rušenie vedení. V mnohých aplikáciách bezdrôtové komunikačné vedenia sú súčasťou nášho každodenného života. Optická bezdrôtová komunikácia je inovatívna technológia, ktorá za posledné tri desaťročia získava čoraz väčšiu pozornosť, pretože rastú nároky užívateľov na rýchlejšiu a bezpečnejšiu komunikáciu.

Tento článok prináša okrem základných informácií o hybridných RF/FSO systémoch aj podrobnejšie informácie typu, ako sú kanálový model alebo hybridná signalizácia týchto systémov. Posledná kapitola je venovaná poveternostným vplyvom na hybridné systémy, ako vypočítať útlm na základe jednotlivých doteraz dostupných modelov a ako nastaviť parametre týchto modelov. Sú dostupné aj rôzne merania, ktoré boli vyhotovené práve na základe tu uvedených modelov. Výsledky týchto meraní však kvôli nedostatku miesta v článku uvedené nie sú. Tieto výsledky sú dostupné pre čitateľov v článkoch uvedených v referenciách.

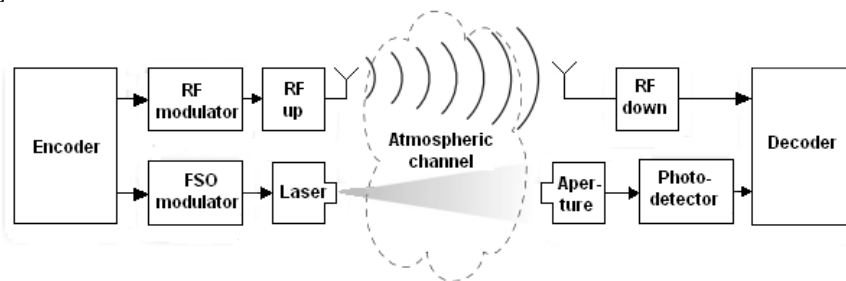
## II. ÚVOD DO HYBRIDNÝCH RF/FSO SYSTÉMOV

V optickej komunikácii voľným prostredím – FSO (Free Space Optic) systémoch na bezdrôtový prenos informácií sa používa optická nosná. FSO systémy poskytujú vysoké dátové rýchlosti, vysokú bezpečnosť prenosu informácií a pracujú v nelicencovaných frekvenčných pásmach. Z nelicencovaných frekvenčných pásiem vyplýva, že nemajú žiadne predpisy ani obmedzenia, čo sa týka frekvencií. Na druhej strane tieto vedenia alebo systémy sú veľmi citlivé a veľmi závislé od atmosférických vplyvov. K týmto atmosférickým vplyvom patrí aj iskrenie spôsobené turbulenciami v atmosfére a spôsobuje náhodné výkyvy prijatého žiarenia optického laserového lúča. Avšak najväčší útlm týchto systémov spôsobuje oblačnosť a hmla. Útlm spôsobený práve týmito atmosférickými vplyvmi je rádovo niekoľko stoviek decibelov na kilometer. Jedným zo spôsobov ako zabezpečiť spoľahlivosť týchto prenosových systémov je skombinovať ich s rádio frekvenčnými – RF (Radio Frequency) systémami. V týchto spojených, čiže hybridných systémoch, keď je FSO časť utlmená oblačnosťou alebo hmlou, RF časť udržuje spoľahlivú komunikáciu aj keď so zníženou prenosovou rýchlosťou. Nevýhodou RF časti týchto systémov je, že sú zase utlmené dažďom, mokrým snehom a iskrením. Z uvedeného vyplýva, že oba kanály sú komplementárne: v FSO časti je signál silne ovplyvnený oblačnosťou a hmlou, čo v RF časti nie je a v RF časti zase je signál ovplyvnený dažďom a snehom. Obidva kanály sú však ovplyvnené iskrením. V súčasnosti existuje veľa prác a teórií, ktoré sa zaoberajú touto problematikou, ale vo väčšine týchto prác sa RF vedenie uvažuje ako záložné vedenie. Je dôležité, aby v hybridných systémoch sa kombinácia FSO a RF systémov prispôbila podmienkam navrhnutého kanála. V doteraz dostupnej literatúre sa dosť málo pojednáva

o vývoji vhodného modelu hybridného RF/FSO kanála, jeho teoretickej analýze a stanovení základných obmedzení komunikácie. V tomto článku preto bude nasledujúca kapitola venovaná kanálovému modelu pre hybridnú RF/FSO komunikáciu. Tento kanálový model je založený na známom paralelnom kanáli, so zohľadnením rozdielnej rýchlosti signalizácie a sú tu zohľadnené útlmové atmosférické javy oboch systémov [1].

### III. KANÁLOVÝ MODEL PRE HYBRIDNÉ RF/FSO SYSTÉMY

Hybridný komunikačný systém uvažovaný v tomto článku je znázornený na Obr. 1. Z obrázku je jasné, že najprv binárna postupnosť prenášanej správy je spoločne zakódovaná do paralelných FSO a RF bitových tokov. Každý bitový tok je modulovaný diskretnou moduláciou, konvertovaný na svoju, respektíve nosnú frekvenciu a prenášaný súčasne cez atmosférický kanál do hybridného prijímača. Na prijímacej strane hybridný prijímač konvertuje smerom dole každý nosný signál, demoduluje a spoločne dekoduje prijatú postupnosť na obnovenie pôvodnej správy. Vývoj takéhoto modelu hybridného systému, ktorý zachytáva všetky základné zložky systému je veľmi ťažký a má zásadný význam. Uplynie niekedy veľa času, kým sa takýto model dostane do praxe. FSO systém pracuje v THz rozsahu frekvencií, kým RF systém pracuje zvyčajne na milimetrových vlnových dĺžkach, t.z. v GHz rozsahu. V dôsledku týchto diametrálne odlišných nosných frekvencií, existuje niekoľko dôležitých fyzikálnych vlastností, ktoré musia byť splnené. Tieto fyzikálne vlastnosti budú spomenuté v pokračovaní tejto kapitoly [1].



Obr. 1 Hybridný RF/FSO komunikačný systém [1]

#### A. Hybridná signalizácia

Vzhľadom k tomu, že nosná frekvencia FSO systémov je niekoľkonásobne vyššia ako nosná frekvencia RF systémov, môže FSO systém podporovať vyššie rýchlosti signalizácie. Táto vlastnosť je vyjadrená cez definíciu o hybridných symboloch. Hybridné symboly pozostávajú z  $N$  FSO symbolov a jedného RF symbolu paralelne. Formálne túto definíciu môžeme zapísať nasledovne [1]:

DEFINÍCIA (Hybridná konštelácia): Nech  $\mathcal{X}_{rf}$  a  $\mathcal{X}_{fso}$  označujú diskkrétne RF a FSO konštelácie, kde  $|\mathcal{X}_{rf}| = M = 2^m$  a  $|\mathcal{X}_{fso}| = Q = 2^q$ . Potom hybridná konštelácia je definovaná ako:

$$\mathcal{X} = \mathcal{X}_{fso}^{\Delta} \times \mathcal{X}_{rf} \quad (1)$$

kde  $N \geq 1$  je kladné celé číslo. Celkovo preto hybridnú konšteláciu môžeme vyjadriť ako  $|\mathcal{X}| = 2^{Nq+m}$ . Pomer medzi počtom FSO bitov a celkovým počtom bitom môžeme vyjadriť ako:

$$\delta = \frac{\Delta Nq}{Nq + m} \quad (2)$$

Pre RF systémy sa bežne moduluje aj amplitúda aj fáza nosnej frekvencie. Preto pre RF systém predpokladáme  $\mathcal{X}_{rf} \subset \mathbb{C}$  je ľubovoľná diskretná komplexná konštelácia, napr. PSK (Phase Shift Keying) alebo QAM (Quadrature Amplitude Modulation). Navyše  $\mathcal{X}_{rf}$  je normalizované

na  $\frac{1}{M} \sum_{x \in \mathcal{X}_{rf}} |x|^2 = 1$ . Aj keď je možné modulovať amplitúdu aj fázu optickej nosnej,

vyžaduje to príslušné heterodýnne alebo homodýnne operácie, ktoré musia byť vykonané pred fotodetekciou. Toto zas vedie k zložitým optickým prijímačom. Preto väčšina FSO systémov používa oveľa jednoduchšiu moduláciu intenzity a priamu detekciu – IM/DD (Intensity Modulation/Direct Detection). To znamená, že je modulovaný len výkon optickej

nosnej. Pre jednoduchosť predpokladajme IM/DD ako  $\mathcal{X}_{fso} \subset R_+^J$ , to znamená, že FSO symboly sú rozdelené do  $J$  časových slotov, pričom cez každý časový slot sa prenáša špecifikované množstvo optického výkonu. Táto všeobecná konštruktívna schéma zahŕňa špeciálne prípady:  $J = 1$  pre pre  $Q$  – árnú PAM (Pulse Amplitude Modulation) alebo  $J = Q$  pre pre  $Q$  – árnú PPM (Pulse Position Modulation). Pre FSO systémy pre všetky  $x = (x[1], \dots, x[J]) \in \mathcal{X}_{fso}$ . Konštruktívna je normalizovaná na  $\sum_{x \in \mathcal{X}} f_X(x) \sum_{j=1}^J x[j] = 1$ , pričom  $f_X(x)$  je diskrétno vstupné rozdelenie pravdepodobnosti [1].

### B. Hybridný kanálový model

Existuje niekoľko dôležitých rozdielov v tom, ako atmosféra ovplyvní obidva typy prenosových systémov. Vplyv atmosférických podmienok na jednotlivé prenosové systémy bude podrobne rozobratý v nasledujúcej kapitole. Už vieme, že obidve nosné frekvencie sú ovplyvnené iskrením, ale s rôznymi časmi koherencie. Z dostupných zdrojov je známe, že optická nosná frekvencia má oveľa kratší koherenčný čas (10s z milisekundy) v porovnaní s RF nosnou frekvenciou (100s milisekundy). Okrem toho každá nosná frekvencia je ovplyvnená časticami porovnateľnými s ich vlnovou dĺžkou. Pri tom všetkom je dôležité poznamenať, že v oboch kanáloch FSO aj RF sú spomenuté útlmové efekty pomalé v porovnaní s ich bežnými signálovými rýchlosťami. Z doterajších poznatkov vieme, že pomalé útlmové kanály sú zjednodušene modelované blokovými útlmovými kanálmi [2], [3]. V tomto nastavení prenášané kódové slovo je rozdelené na konečný počet blokov, z ktorých každý pozostáva z rovnakého počtu symbolov. Tento model blokových útlmových kanálov je vhodný ako pre FSO tak aj pre RF kanály. Avšak treba brať do úvahy rozdielny koherenčný čas medzi dvomi kanálmi. Aby sa toto zabezpečilo, treba splňať nasledujúci predpoklad: FSO systému prislúcha kódové slovo  $A$  s dĺžkou bloku  $K$  a RF systému zase kódové slovo  $B$  s dĺžkou bloku  $L$ , pričom  $A \geq B$  a  $AK = BL$ . S týmito predpokladmi pre paralelne FSO a RF blokové útlmové kanály platí:

$$\text{Pre FSO systém} \quad r_a[k, n] = \rho h_a \frac{p_a}{N} x_a[k, n] + \lambda \quad (3)$$

kde:  $r_a[k, n] = (r_a[k, n, 1], \dots, r_a[k, n, J]) \in R_+^J$  označuje  $J$  vektor predstavujúci prijatú optickú energiu v každom z  $J$  časových slotov FSO symbolu  $n = 1, \dots, N$  hybridného symbolu  $k = 1, \dots, K$  v bloku  $a = 1, \dots, A$ ;  $x_a[k, n] = (x_a[k, n, 1], \dots, x_a[k, n, J]) \in \mathcal{X}_{fso}$  označuje  $n$  – tý prenesený FSO symbolový vektor hybridného symbolu  $k$  v bloku  $a$ ;  $\rho$  modeluje dlho trvajúci útlmový efekt;  $h_a > 0$  označuje útlm v bloku  $a$  v dôsledku iskrenia.  $p_a$  označuje výkon pridelený bloku  $a$  a  $\lambda \geq 0$  modeluje žiarenie na pozadí (predpokladá sa, že je konštantné). Faktor  $1/N$  zabezpečuje, aby komponenty FSO a RF hybridného symbolu mali rovnakú energiu.

Pre RF kanál prijatý  $l$  – tý symbol bloku  $b$  môžeme vyjadriť podobne:

$$\hat{y}_b[l] = \sqrt{\hat{\rho} \hat{h}_b p_b} \hat{x}_b[l] + \hat{\sigma} \hat{z}_b[l] \quad (4)$$

pre  $l = 1, \dots, L$  a  $b = 1, \dots, B$ , kde:  $\hat{x}_b[l] \in \mathcal{X}_{rf}$  označuje prenesený symbol;  $\hat{\rho}$  modeluje dlho trvajúci útlmový efekt;  $\hat{h}_b > 0$  označuje útlm v bloku  $b$  v dôsledku iskrenia.  $\hat{p}_b$  označuje výkon pridelený bloku  $b$ ;  $\hat{z}_b[l] \approx N_c(0, 1)$  modeluje nulovú strednú hodnotu prídavného Gaussovho bieleho šumu – AWGN (Additive White Gaussian Noise).  $\hat{\sigma}^2$  označuje RF šumový výkon.  $\hat{h}_a$  a  $\hat{h}_b$  sú nezávislé.

Parametre  $0 < \rho < 1$  a  $\hat{\rho} < 1$  modelujú rozdiely týchto dvoch paralelných kanálov, ako sú napr. útlmové efekty vplyvom dažďa, hmly alebo oblačnosti. Takisto modelujú aj ďalšie parametre ako sú clona/zisk antény alebo strata počas šírenia sa signálu. Ak  $\rho > \hat{\rho}$  FSO kanál je oveľa



silnejší ako RF kanál, to znamená, že takýmto spôsobom sa napr. modeluje útlm vplyvom dažďa. Na druhej strane ak je  $\rho \langle \hat{\rho}$  RF kanál je oveľa silnejší ako FSO kanál. Týmto spôsobom sa dá modelovať útlm v dôsledku hmly. V praxi však sa  $\rho$  a  $\hat{\rho}$  náhodne menia v čase [1].

#### IV. VPLYV POČASIA NA HYBRIDNÉ RF/FSO SYSTÉMY

Ako už bolo spomenuté vyššie táto kapitola je venovaná vplyvom atmosférických podmienok na hybridné RF/FSO systémy. Pre FSO je veľmi dôležitá priama viditeľnosť, čo je niekedy veľmi ťažké zabezpečiť, najmä, keď je vonku hustá hmla. Hmla dokáže spôsobiť útlm signálu od 120 dB/km až do 480 dB/km, všetko závisí od lokality a od ročného obdobia. Preto pri hybridných systémoch sa používa RF systém ako záložný, pri ktorých útlm v dôsledku hmly je zanedbateľný. Táto kapitola je rozdelená na štyri podkapitoly, kde sú podrobne opísané jednotlivé typy útlmov na tieto dva typy systémov.

##### A. Útlm spôsobený hmlou

Najväčším problémom pri realizácii vonkajších optických bezdrôtových vedení je atmosférický útlm spôsobený absorpciou a rozptylom. Absorpcia optických signálov je vo veľkej miere zapríčinená čistočkami vody a oxidu uhličitého. Rozptyl optického signálu je zase spôsobený hmlou, dažďom, snehom a oblačnosťou. Takýto rozptyl spôsobuje, že časť svetelného lúča putujúceho smerom od zdroja sa odkloní od prijímača. Zo všetkých tu spomenutých atmosférických javov pre FSO systémy je najhoršia hmla. Hmla spôsobuje značný útlm optických signálov na dlhšiu dobu a preto je hlavnou prekážkou pri dosahovaní požadovanej dostupnosti FSO systémov. Hlavným dôvodom vysokého útlmu v dôsledku rozptylu pre rôzne formy hmly je to, že veľkosť hmlových čistočiek je porovnateľná s prenosovými vlnovými dĺžkami optických a blízkych infračervených vln. Tento rozptylový jav je nazývaný ako Mieho rozptyl a optický útlm v prípade kvapôčok hmly je možné presne predpovedať na základe teórie Mieho rozptylu. Avšak na toto sú potrebné zložité výpočty a je potrebná aj podrobná informácia o parametroch hmly, ako sú veľkosť častíc a ich rozdelenie, index lomu a podobne, ktoré nemusia byť k dispozícii na konkrétnom mieste. Iným spôsobom ako je možné predpovedať útlm spôsobený hmlou je na základe informácií o rozsahu viditeľnosti. Rozsah viditeľnosti je vzdialenosť k objektu, kde odraz obrazu kvapiek je 5% z rozsahu, ak by bol objekt vedľa. Vlnová dĺžka 550 nm sa bežne používa na meranie rozsahu viditeľnosti a je tiež akceptovaná ako referenčná vlnová dĺžka pre meranie rozsahu viditeľnosti. Na tejto vlnovej dĺžke sa rozsah viditeľnosti bežne meria na letiskách a v meteorologických stanicách. Na predpovedanie útlmov na základe viditeľnosti sa bežne používajú tri modely [8]. Tými modelmi sú: Kruse, Kim a Al-Naboulsi modely [4] – [7]. Útlmový koeficient pre modely Kruse a Kim je daný podľa (5):

$$\alpha_{Fog} = \frac{10 \log V\%}{V(km)} \left( \frac{\lambda}{\lambda_0} \right)^{-q} \quad (5)$$

kde  $V(km)$  je rozsah viditeľnosti,  $V\%$  je prenos vzduchových čistočiek v percentách pri jasnej oblohe,  $\lambda$  v nm je vlnová dĺžka,  $\lambda_0$  je referenčný rozsah viditeľnosti (550 nm) a  $q$  je rozdelenie veľkosti koeficientu rozptylu. Pre Krusov model [4], parameter  $q$  je:

$$q = \begin{cases} 1,6 & \dots \dots \dots \text{ak } V > 50km \\ 1,3 & \dots \dots \dots \text{ak } 6km < V < 50km \\ 0,585V^{\frac{1}{3}} & \dots \dots \dots \text{ak } V < 6km \end{cases} \quad (6)$$

Pre Kimov model [5], parameter  $q$  je:

$$q = \begin{cases} 1,6 & \dots \dots \dots \text{ak } V > 50km \\ 1,3 & \dots \dots \dots \text{ak } 6km < V < 50km \\ 0,16V + 1,34 & \text{ak } 1km < V < 6km \\ V - 0,5 & \dots \dots \dots \text{ak } 0,5km < V < 1km \\ 0 & \dots \dots \dots \text{ak } V < 0,5km \end{cases} \quad (7)$$

Zo (6) vyplýva, že pre všetky poveternostné podmienky bude menší útlm na vyšších vlnových dĺžkach.

Al-Naboulsi [6], [7] navrhol dva rôzne modely, prvý je pre advekčnú hmlu a druhý pre radiačnú hmlu. Advekčná hmla je tvorená pohybom vlhkých a teplých čistočiek vzduchu



nad chladnejšie pobrežné povrchy. Radiačná hmľa zase sa vzťahuje k zemi chladenej radiáciou cez kontinentálny povrch. Pre adveknú hmľu, teda platí vzťah:

$$\alpha_{ADV}(\lambda) = \frac{0,11478\lambda + 3,8367}{V} \quad (8)$$

A pre radiačnú hmľu platí:

$$\alpha_{RAD}(\lambda) = \frac{0,18126\lambda^2 + 0,13709\lambda + 3,7502}{V} \quad (9)$$

Konkrétny útlm pre obidva typy hmíel môže byť vypočítaný nasledovným spôsobom:

$$A_{FOG}(dB/km) = \frac{10}{\ln(10)}(\alpha(\lambda)) \quad (10)$$

### B. Vplyv hmly na GHz vedenia

Malé kvapôčky obsiahnuté v hmle spôsobujú rozptyl a významný útlm pre záložné rádio frekvencie, väčšie ako 10 GHz, zatiaľ čo vplyv na frekvenciách nižších ako 10 GHz je zanedbateľný. Pre kvapôčky hmly s rozmermi menšími ako 0,01 cm a frekvencie nižšie ako 200 GHz platí Rayleigha aproximácia. Na vyjadrenie útlmu je možné použiť celkový obsah vody na jednotku objemu. Útlm je potom možné vyjadriť nasledovne:

$$\gamma_{FOG} = K_l M \quad (11)$$

kde:  $\gamma_{FOG}$  je útlm v dB/km v oblaku alebo hmle,  $K_l$  je útlmový koeficient v (dB/km)/(g/m<sup>3</sup>) a  $M$  je hustota vody v oblaku alebo v hmle v (g/m<sup>3</sup>) [8].

Útlmový koeficient sa dá vypočítať použitím matematického modelu založeného na Rayleigh rozptyle. Tento model používa dvojitý Debyeov model pre dielektrickú permitivitu  $\mathcal{E}(f)$  vody, kde  $\mathcal{E}(f) = \mathcal{E}' + j\mathcal{E}''$ . Útlmový koeficient je potom daný ako:

$$K_l = \frac{0,819f}{\mathcal{E}''(1 + \eta^2)} \quad (12)$$

kde:  $f$  je frekvencia v GHz,  $\mathcal{E}$  je komplexná dielektrická permitivita a  $\eta = \frac{2 + \mathcal{E}'}{\mathcal{E}''}$  [8].

Komplexná dielektrická permitivita vody je vyjadrená ako:

$$\mathcal{E}''(f) = \frac{f(\epsilon_0 - \epsilon_1)}{f_p \left(1 + \left(\frac{f}{f_p}\right)^2\right)} + \frac{f(\epsilon_1 - \epsilon_2)}{f_s \left(1 + \left(\frac{f}{f_s}\right)^2\right)} \quad (13)$$

$$\mathcal{E}'(f) = \frac{f(\epsilon_0 - \epsilon_1)}{\left(1 + \left(\frac{f}{f_p}\right)^2\right)} + \frac{f(\epsilon_1 - \epsilon_2)}{\left(1 + \left(\frac{f}{f_s}\right)^2\right)} + \epsilon_2 \quad (14)$$

kde:  $f_p$  a  $f_s$  sú hlavné a vedľajšie frekvencie v GHz a sú vyjadrené ako

$$f_p = 20,09 - 142 \left(\frac{300}{T} - 1\right) + 294 \left(\frac{300}{T} - 1\right)^2, \quad f_s = 590 - 1500 \left(\frac{300}{T} - 1\right), \quad T \text{ je teplota}$$

v Kelvinoch a  $\epsilon_0 = 77,6 - 103,3 \left(\frac{300}{T} - 1\right)$ ,  $\epsilon_1 = 5,48$ ,  $\epsilon_2 = 3,1$  [8].

### C. Útlm spôsobený dažďom

Na signál prechádzajúci atmosférou pôsobí okrem hmly aj dažď. Ak je rozmer dažďových kvapiek veľký, vzniká odraz a rozptyl nezávislý od vlnovej dĺžky. Útlm sa lineárne zvyšuje so silou a intenzitou dažďa. Útlm vzhľadom k intenzite dažďa  $R$  (mm/h) pre optické bezdrôtové vedenia je daný nasledovne:

$$\alpha_{Rain} = 1,076R^{0,67} \quad (15)$$

Najvšeobecnejšia forma distribučnej funkcie  $N_a$  veľkosti kvapiek je podľa Weibulla [9]:

$$N_a = \frac{N_T}{a_0 V_a} \phi n(n) \left(\frac{a}{a_0}\right)^{n-1} e^{-\phi(n) \frac{a}{a_0}} \quad (16)$$

kde:  $a_0 = d(z_a)^b e^{-cz_a}$  a  $\phi(n) = \Gamma^n \left(1 + \frac{1}{n}\right)$ .  $z_a$  je intenzita dažďových kvapiek s polomerom

$a$ ,  $d = 0,941$ ,  $b = 0,336$ ,  $c = 0,471102$  a  $n = 3$ .  $N_T$  reprezentuje celkový počet kvapiek rôznych veľkostí na jednotku objemu. Celkový rozptylový koeficient vplyvom dažďa môže byť vyjadrený ako:

$$\beta_{scat}^{rain} = \sum_a \pi a^2 N_a Q_{scat} \left(\frac{a}{\lambda}\right) \quad (17)$$

kde  $Q_{scat}$  je účinnosť rozptylu. Aj keď šírenie mikrovlnových signálov je značne narušené hmlou, oblačnosťou a prachovými častočkami, hlavným útlmovým faktorom frekvencií nad 10 GHz je dážď [8]. Existujú dva populárne modely pre predikciu útlmu vplyvom dažďa: ITU-R terestriálny model, bližšie popísaný v [10] a Craneov model, bližšie popísaný v [11].

#### Útlm spôsobený snehom

Útlm hybridných systémov môže spôsobiť mokrý aj suchý sneh. Ak  $S$  je intenzita sneženia v mm/h, potom útlm  $\alpha_{Snow}$  v dB/km je daný nasledovne:

$$\alpha_{Snow} = aS^b \quad (18)$$

Ak  $\lambda$  je vlnová dĺžka, potom pre suchý sneh platia tieto hodnoty parametrov:  $a = 5,42 \times 10^{-5} \lambda + 5,4958776$  a  $b = 1,38$ . Pre mokrý sneh platia tieto hodnoty parametrov:  $a = 1,023 \times 10^{-4} \lambda + 3,7855466$  a  $b = 0,72$ .

Útlm rádiových vln vplyvom snehu je ťažké analyzovať, vzhľadom k fyzickej povahe snehových vločiek. Existuje totiž veľa odlišností v tvare, type a veľkosti snehových vločiek [8].

#### V. ZÁVER

V súčasnej uponáhľanej dobe je komunikácia a výmena informácií jedna z najdôležitejších úloh. Optické siete predstavujú pomerne kvalitný, stabilný a spoľahlivý prenos informácií. Napriek svojim výhodám tieto bezdrôtové systémy majú aj svoje nevýhody, kde patrí najmä priama viditeľnosť. Existuje veľa dostupnej literatúry, ktorá sa zaoberá touto problematikou, keďže je to veľmi populárny a rozvíjajúci sa spôsob komunikácie.

#### POĎAKOVANIE

Táto práca bola podporovaná Kultúrnou a edukačnou grantovou agentúrou (KEGA) Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky na základe projektu č. „063TUKE-4/2013 - Použitie diaľkovo ovládaného optického vláknového refraktometra vo výučbe“ a projektu č. „006TUKE-4/2014 - Experimentálny model PON siete TUKE a jeho použitie vo výučbe“. Táto práca je tiež výsledkom zmluvy „APVV-0025-12 - Predchádzanie vplyvu stochastických mechanizmov vo vysokorýchlostných plne optických sieťach“ podporovanej Agentúrou na podporu výskumu a vývoja (APVV) Slovenskej republiky.

#### REFERENCIE

- [1] S. Arnon, J. R. Barry, G.K. Karagiannidis, R. Schober, M. Uysal, “Advanced Optical Wireless Communication Systems,” Cambridge University Press, 2012.
- [2] L. H. Ozarow, S. Shamai, A. D. Wyner, “Information theoretic considerations for cellular mobile radio”, IEEE Trans. Veh. Technol., vol. 43, no. 2, pp. 359 - 378, May 1994
- [3] E. Biglieri, J. Proakis, S. Shamai, “Fading channels: information – theoretic and communications aspects”, IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2619 – 2692, Oct. 1998
- [4] P. W. Kruse, “Elements of Infrared Technology: Generation, Transmission and Detection”, J. Wiley and Sons, New York 1962
- [5] I. Kim, B. McArthur, E. Korevaar, “Comparison of Laser Beam Propagation at 785 and 1550 nm in Fog and Haze for Optical Wireless Communications”, Proc. SPIE, Vol. 4214, pp. 26 – 37, 2001
- [6] M. Al Naboulsi, H. Sizun, F. de Fornel, “Fog Attenuation Prediction for Optical and Infrared Waves”, Optical Eng., 42(2), pp. 319 – 329, February 2004
- [7] O. Bouchet, T. Marquis, M. Chabane, M. Al Naboulsi, H. Sizun, “FSO and Quality of Service Software Prediction”, Proc. SPIE, Vol. 5892, pp. 01 – 12, 2005
- [8] F. Nadeem, V. Kvicera, M. S. Awan, E. Leitgeb, S. S. Muhammad, G. Kandus, “Weather Effects on Hybrid FSO/RF Communication Link”, IEEE Journal on Selected Areas in Communications, Vol. 27, No. 9, pp. 1687 – 1697, December 2009
- [9] G. Brussaard, P. A. Watson, “Atmospheric Modelling and Millimetre Wave Propagation”, ISBN: 978-0-412 -56230-3, 1994
- [10] ITU-R, P.838-1: “Specific Attenuation Model for Rain for use in Prediction Methods”
- [11] R. K. Crane, “Prediction of Attenuation by Rain”, IEEE Trans. Commun., COM-28, pp. 1717-1733, 1980

# Vyhľadavanie podozrivých tém v sociálnom webe prostriedkami strojového učenia

<sup>1</sup>Libuša VANICKÁ, <sup>2</sup>Kristína MACHOVÁ

<sup>1</sup> Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>libusa.vanicka@gmail.com, <sup>2</sup>kristina.machova@tuke.sk

**Abstrakt** — Článok sa zaoberá klasifikovaním nových príspevkov do tried na základe konverzačného obsahu. Prvá časť opisuje princíp a základy metód strojového učenia. Bližšie je opísaný aj princíp fungovania použitého rozhodovacieho stromu C4.5, ktorý tvorí dôležitú súčasť riešenia úlohy. Druhá časť práce sa zaoberá opisom metódy modelovania témy za účelom identifikácie podozrivého konverzačného obsahu. Taktiež je opísaný samotný proces klasifikácie daného príspevku k podozrivej téme použitím rozhodovacieho stromu C4.5. Záver článku obsahuje postup riešenia a samotné experimenty rôznych príkladov s dosiahnutými výsledkami.

**KLúčové slová** — modelovanie témy, strojové učenie, identifikácia podozrivého obsahu, rozhodovací strom C4.5, konverzačný obsah

## I. ÚVOD

V dnešnej technologicky vyspelejšej dobe, takmer každý z nás žije nielen svojim sociálnym životom, ale aj aktívnym životom na sociálnom webe. Avšak so stúpajúcim záujmom o sociálne siete stúpa aj riziko, resp. hrozby zo strany podozrivých aktérov. Nebezpečenstvo sa nachádza rovnako vo verejných diskusiách, kde ktokoľvek, mnohokrát aj anonym, môže prispieť svojim netolerantným názorom k danej téme. Avšak mnohokrát tieto príspevky môžu obsahovať skrytý podtón, viac, či menej osočujúci iných. Z tohto dôvodu sa kladie čoraz väčší dôraz na bezpečnosť na sociálnych sieťach. Hlavne, aby sa internet, a tým aj diskusie stali bezpečnejším miestom aj pre deti.

Existuje mnoho hrozieb, ktoré je potrebné odhaliť. Tento článok sa zameriava na odhalenie príspevkov s rasistickým náznakom. Aj keď každým dňom sú vyvíjané nové technológie, ktoré by tieto príspevky, resp. samotných prispievateľov odhalili, stále existuje široké spektrum neobjavených možností ako uchrániť diskusné fóra od podozrivých aktérov. Takéto odhalenie je často náročné. Vhodným prístupom pre riešenie tohto problému by mohli byť metódy strojového učenia [1].

## II. ROZHODOVACIE STROMY

Rozhodovacie stromy predstavujú snáď najznámejšiu klasifikačnú metódu strojového učenia. Hlavným dôvodom rozšíreného používania rozhodovacích stromov v súčasnosti je to, že reprezentujú naučenú znalosť spôsobom, ktorý je dobre pochopiteľný a čitateľný aj pre laika.

Rozhodovacie stromy klasifikujú inštanície postupným delením priestoru inštančiami smerom nadol v strome od koreňového uzla k listovým uzlom. Listový uzol poskytuje informáciu potrebnú na klasifikáciu inštanície. Každý uzol v strome, okrem listových, špecifikuje test atribútu danej inštanície a každá vetva zostupne od tohto uzlu korešponduje s jednou z možných hodnôt pre tento atribút. Výhody použitia metód rozhodovacích stromov sú nasledovné:

- Jednoduché na pochopenie a interpretáciu.
- Dokážu pracovať s veľkými dátovými množinami.
- Sú robustné.
- Dokážu generovať zrozumiteľné pravidlá.
- Dokážu vykonávať klasifikáciu bez nutnosti zdĺhavých výpočtov.
- Dokážu poskytnúť jasnú indikáciu, ktoré polia sú najdôležitejšie pre klasifikáciu.

- Sú schopné pracovať so spojitými aj diskretnými dátami.
- Nevýhody použitia metódy rozhodovacích stromov:
- Rozhodovacie stromy sú náchylnejšie k chybám v klasifikačných problémoch s vyšším počtom tried a relatívne malým počtom tréningových príkladov.
  - V každom uzle každý atribút rozdelenia podľa inštancií musí byť zaradený pred nájdením jeho najlepšieho delenia. Pri niektorých algoritmoch sa používajú kombinácie polí a hľadanie musí byť tvorený pri optimálnom kombinovaní, resp. nastavení váh. Algoritmy zaoberajúce sa aj orezávaním preučeného stromu sú výpočtovo náročné.

#### A. Rozhodovací strom C4.5

C4.5 je algoritmus, ktorý navrhol Ross Quinlan v roku 1993 a používa sa na generovanie rozhodovacích stromov. C4.5 je vylepšeným a doplneným algoritmom ID3, ktorý bol rovnako navrhnutý Rossom Quinlanom. Rozhodovacie stromy vygenerované pomocou algoritmu C4.5 sú používané na klasifikačné úlohy. Klasifikátorom sa rozumie algoritmus, ktorý implementuje triedenie. Niekedy sa tým odkazuje na matematickú funkciu, realizovanú klasifikačným algoritmom, ktorý mapuje vstupné dáta do tried. V porovnaní s algoritmom ID3 pracuje so *spojitými* - reálnymi, rovnako ako aj *diskretnými* - nominálnymi atribútmi. V prípade, že dáta, ktoré chceme použiť pri rozhodovacích stromoch nie sú typovo vhodné, je nutné pretransformovať tieto dáta na nominálne alebo reálne atribúty. Tento algoritmus sa radí medzi neinkrementálne algoritmy a vygenerovaný strom je budovaný zhora nadol [2].

### III. PREDSPRACOVANIE DÁT

Dôležitou súčasťou tejto práce bolo okrem iného aj predspracovanie, ktoré sa skladalo z viacerých krokov:

1. **Oprava gramatických chýb** – Bolo potrebné overiť gramatickú správnosť dát, ktoré mali slúžiť ako tréningová množina.
2. **Odstránenie interpunkčných znamienok** - dôvodom bolo, že väčšina príspevkov v diskusných fórach je napísaných bez použitia interpunkcie a stáva sa, že autor použije nesprávnu interpunkciu. Potom by bolo veľmi náročné pracovať s takouto vzorkou tréningových príkladov.
3. **Odstránenie stopových slov** - stopové slová predstavujú slová bez samostatného významu a teda nie sú relevantné pri riešení danej problematiky. Tieto stopové slová sa odstránia počas behu programu. Všetky stopové slová s nachádzajú v textovom súbore *stop-words.txt*.

#### A. Určenie váh slovám a príspevkom

Pri spracovávaní textu je potrebné identifikovať slová, ktoré sú dôležité pri reprezentácii textu. Aby bolo možné tieto slová rozlíšiť, definuje sa relatívna hodnota, tzv. váha každého slova. Váha určuje selektívnu silu slova, teda dôležitosť termu pre daný dokument vzhľadom ku celému korpusu dokumentov. *Váhovaním* sa rozumie proces definovania a výpočtu váh.

Odhalenie príspevkov s rasistickým motívom je náročné z toho dôvodu, že tieto príspevky sa na prvý pohľad nelíšia od bežných príspevkov a teda nie je možné priamo identifikovať slová, ktoré sú stopercentne rasistické. Z tohto dôvodu bolo potrebné zvoliť metódu, pomocou ktorej by bolo možné každému slovu, resp. príspevku určiť jeho závažnosť pre zvolenú tému. Vychádzalo sa zo známej váhovacej techniky TF-IDF [3].

#### B. Proces váhovania

Proces váhovania prebiehal nasledovne. Najprv sa určili váhy jednotlivým slovám každého príspevku. Ďalej boli určené váhy jednotlivých príspevkov vypočítaním aritmetického priemeru váh jednotlivých slov v danom príspevku. Následne týchto 100 priemerných hodnôt, ktoré sa nachádzali v intervale  $< 0,04387; 2 >$ , bolo diskretizovaných na 4 menšie intervaly, ktoré reprezentovali triedu závažnosti a to nasledovným spôsobom:

- $< 0 ; 0,25 >$  - nahradené triedou 1
- $(0,25 ; 0,5 >$  - nahradené triedou 2
- $( 0,5 ; 0,7 >$  - nahradené triedou 3
- viac ako 0,7 - nahradené triedou 4

Nakoľko sa tie isté slová opakovali vo viackrát, bol vytvorený textový súbor *priemery.txt*, kde sa každé slovo nachádza už iba raz a jeho váha je aritmetickým priemerom všetkých váh daného slova v pôvodnej množine príspevkov. Tento súbor tvorí 394 slov s ich váhami. Slová v súbore *priemery.txt* sú zoradené od najkratšieho po najdlhšie a pod každým slovom, čiže v nasledujúcom riadku sa nachádza priemerná váha.

Avšak samotné váhovanie pri riešení tohto problému nestačí nakoľko môžu nastať prípady, kedy dané slovo má samotnú váhu malú, ale z hľadiska miery podozrivosti slova je významné a teda jeho váha by mala byť zvýšená. Tento problém bol riešený vytvorením textového súboru s názvom *tabu.txt*, ktorý obsahuje slová z morálneho hľadiska neprípustné alebo považované za urážlivé.

Podľa toho koľko takýchto slov sa nachádza v danom príspevku priradí sa príspevku trieda nasledovným spôsobom:

- ak sa v príspevku nachádza 0 alebo 1 tabu slovo, príspevku sa priradí trieda 1
- ak príspevok obsahuje 2 slová, príspevku je priradená trieda 2
- ak sa v príspevku nachádzajú 3 tabu slová, príspevku je priradená trieda 3
- nakoniec, ak je v príspevku viac ako 3 slová, teda 4 a viac, trieda je 4.

Vďaka týmto podmienkam sa zvyšuje presnosť určenia stupňa miery rasizmu, ktorý príspevok môže obsahovať.

Ďalším krokom k vytvoreniu konečnej trénovacej množiny je určenie tried, do ktorých sa následne budú klasifikovať nové príklady. Na základe triedy danej pri váhovaní a triedy danej pri vyhodnotení tabu slov. Výsledné triedy, reprezentujúce stupeň rasizmu sú interpretované nasledovne:

- Trieda 1 - *malý náznak rasizmu*
- Trieda 2 - *mierny náznak rasizmu*
- Trieda 3 - *stredný náznak rasizmu*
- Trieda 4 - *veľký náznak rasizmu*

Samotné trénovanie prebiehalo na základe trojnásobnej validácie.

#### IV. IMPLEMENTÁCIA RIEŠENIA

Implementácia navrhutej metódy pracuje v nasledovných krokoch:

1. Vloženie príspevku, ktorý chceme otestovať - je potrebné zadávať príspevky bez interpunkcie a bez znaku na konci príspevku.
2. Úprava dát - program je nastavený tak, že po vložení príspevku sa odstránia, resp. ignorujú diakritické znamienka, interpunkcia a odstránenie stop slov. Vďaka tejto počiatkovej úprave je nasledovné váhovanie presnejšie a s väčšou výpovednou hodnotou.
3. Vyhodnotenie príspevku - výstupom v programe sú hodnoty, resp. váhy príspevku, ako aj dodatočné informácie o zadanom príspevku, napr. počet slov, počet odstránených stop slov, počet tabu slov a textový súbor, ktorý obsahuje dané triedy príspevku vo formáte arff pre spustenie v prostredí WEKA.
4. Aplikovanie metódy rozhodovacieho stromu C4.5 v prostredí WEKA - na naučenej trénovacej množiny sa tu otestuje ľubovoľný nový príspevok do diskusie pomocou algoritmu, ktorý generuje rozhodovací strom C4.5. Výsledkom je konečné zaradenie príspevku do triedy rasistických príspevkov, resp. nezaradenie.

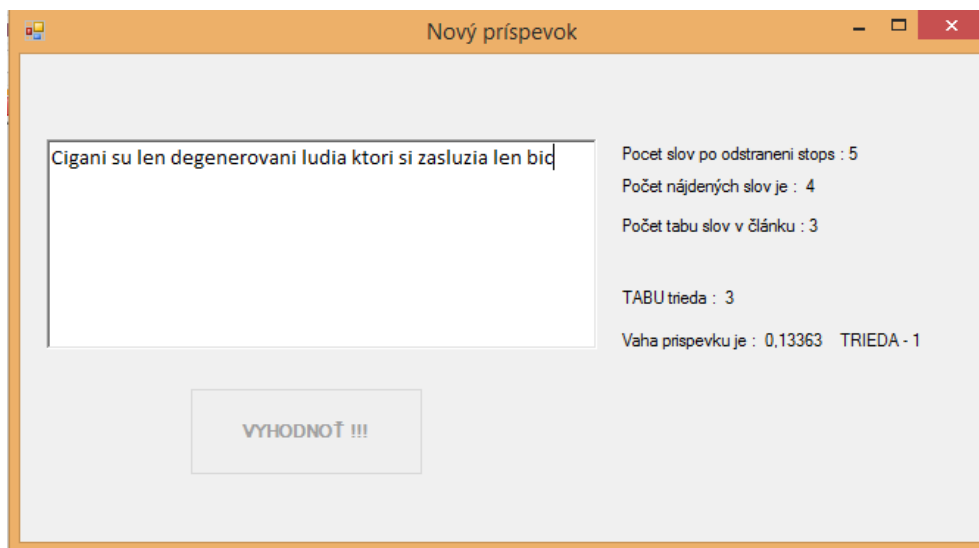
#### V. EXPERIMENTY

Experiment v tejto časti je zameraný na automatickú klasifikáciu nového diskusného príspevku do rasistickej triedy (resp. nie) a porovnanie tohto záveru so záverom experta na bezpečnosť webu.

##### A. Experiment 1

Ako prvý experiment bol zvolený nový príspevok určený na klasifikáciu v znení: "*Cigani su len degenerovani ludia ktori si zaslužia len bic*".

Výstup z programu je zobrazený na Obr.1.



Obr.1. Klasifikácia implementácie v prvom experimente.

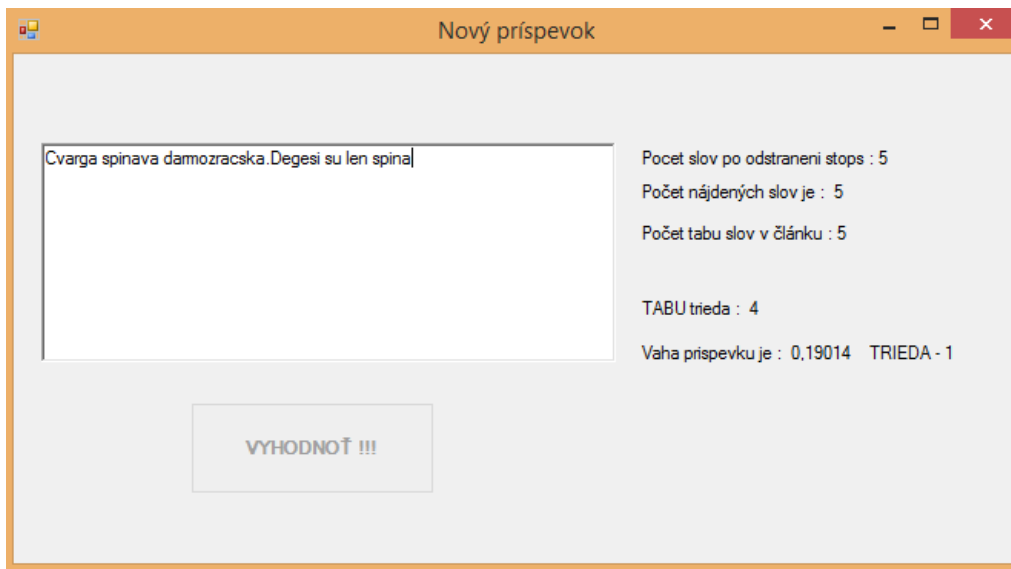
Po zadaní výstupného textového súboru do programu WEKA boli pri trojitej validácii dosiahnuté nasledovné výsledky:

- 1. prípad - trieda 2
- 2. prípad - trieda 2
- 3. prípad - trieda 1

Ako je znázornené vyššie v dvoch prípadoch bol nový príklad klasifikovaný do triedy 2 a v poslednom prípade do triedy 1. Dôvodom tohto rozdielu môže byť to, že v tretej trénovacej množine sa nachádzal hraničný príklad, čiže príklad bol klasifikovaný do inej triedy ako mal byť, alebo v trénovacej množine bolo viacero príkladov klasifikovaných nesprávne.

### B. Experiment 2

Nový testovaný príspevok je v znení: "*Cvarga spinava damozracska.Degesi su len spina*". Výstupom z programu pre daný príspevok je zobrazený na Obr.2.



Obr.2. Klasifikácia implementácie v druhom experimente

Výsledkom testovania v programe WEKA je daný príspevok zaradený do týchto tried:

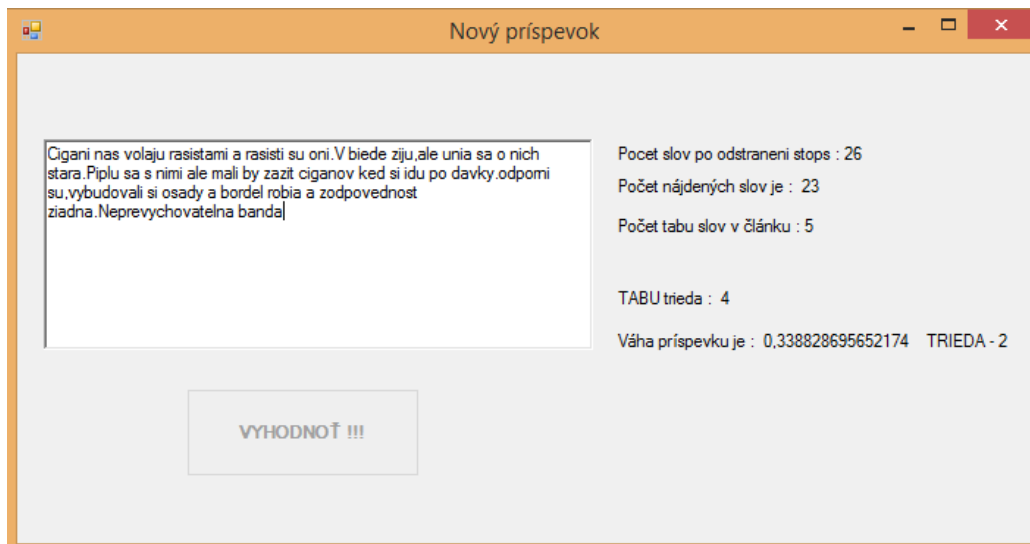
- prípad - trieda 3
- prípad - trieda 3
- prípad - trieda 3

Zhoda je v tomto experimente stopercentná, čiže môžeme povedať, že príspevok jednoznačne patrí do triedy 3. Môžeme teda zhrnúť, že všetky tri časti trojitej validácie v tomto prípade boli natrénované s dostatočnou presnosťou vďaka čomu aj nový príklad klasifikovali jednoznačne.

### C. Experiment 3

V prípade tretieho experimentu bol použitý dlhší príspevok, aby bolo jasné ako sa implementácia bude chovať pri vyšších váhach príspevkov. Bol uvažovaný príspevok v znení: "Cigani nas volaju rasistami a rasisti su oni.V biede ziju,ale unia sa o nich stara.Piplu sa s nimi ale mali by zazit ciganov ked si idu po davky.odporni su,vybudovali si osady a bordel robia a zodpovednost ziadna.Neprevychovatelna banda "

Výstup programu je v podobe Obr.3.



Obr.3 Klasifikácia implementácie v treťom experimente

Výsledkom testovania v programe WEKA je zaradenie daného príspevku do nasledovných tried:

- prípad - trieda 3
- prípad - trieda 3
- prípad - trieda 3

Aj v tomto experimente bol príklad klasifikovaný jednoznačne do triedy 3.

## VI. ZÁVER

Cieľom tohto príspevku bolo navrhnúť vhodnú metódu strojového učenia na modelovanie tém rasistického charakteru v rámci konverzačného obsahu. Bol navrhnutý a implementovaný prístup, založený na váhovacej technike TF-IDF a rozhodovacích stromoch C4.5, na identifikáciu nevhodných, rasisticky mierených príspevkov. Algoritmus C4.5 bral do úvahy dva tribúty – triedu, ktorá bola určená váhovaním a tabu triedu, ktorá bola identifikovaná pomocou tabu slovníka. Na základe týchto dvoch výsledkov prebiehalo konečné klasifikovanie do tried, ktoré reprezentovali silu rasistického náznaku v danom príspevku.

Tento prístup sa ukázal ako vhodný na riešenie takéhoto typu úloh. Implementácia fungovala rýchlo, spoľahlivo a efektívne. Neboli zaznamenané žiadne chybové hlásenia. Výhodou bola aj rýchlosť výpočtov.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] L. Vanická, „Vyhľadavanie podozrivých aktérov v sociálnom webe využitím metód strojového učenia“, Diplomová práca, Technická univerzita, Košice, 2014, 100s.
- [2] J.R. Quinlan, “Bagging, boosting and C4.5.” In Proc. of the Fourteenth National Conference on Artificial Intelligence, 1996.
- [3] G. Salton, C. Buckley, “Term weighting approaches in automatic text retrieval”. Information Processing and Management, 24(5), 1988, 513-523.

# Detekcia typu prehovoru používateľa

<sup>1</sup>Stanislav Ondáš, <sup>1</sup>Maroš Minda

<sup>1</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulte Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>stanislav.ondas@tuke.sk,

**Abstrakt** — Článok sa zaoberá detekciou typu prehovoru používateľa, ktorá súvisí s interakčnými zámermi používateľa v komunikačnom procese. Postojová modálnosť prehovoru je zakódovaná v prozodických a syntaktických vlastnostiach prehovoru. Hlavným prozodickým prvkom je melódia vety. Predkladaný článok popisuje jednoduchý systém pre detekciu typu prehovoru používateľa. Ako parametrický popis je použitý vektor hodnôt frekvencie základného tónu, meraného na znelých úsekoch prehovoru. Pre úlohu detekcie sme použili algoritmus DTW, používaný aj na automatické rozpoznávanie izolovaných slov. Úspešnosť popisovaného systému bola vyhodnotená na novo-vytvorenej databáze vetných typov.

**Kľúčové slová** — kadencia, postojová modálnosť, frekvencia základného tónu, DTW

## I. ÚVOD

Ak hovoríme o detekcii typu vety, máme na mysli jej modálnu stavbu, konkrétnejšie postojovú modálnosť. Postojová modálnosť vyjadruje komunikatívny zámer hovoriaceho [1]. Vety podľa postojovej modálnosti delíme na oznamovacie, opytovacie, rozkazovacie a želacie.

Detekcia typu vety je dôležitou súčasťou moderných komunikačných systémov určených pre komunikáciu človeka so strojmi prirodzenou rečou vo forme dialógu. Detekcia typu vety je potrebná pre mechanizmy riadenia dialógu medzi používateľom a systémom, najmä v prípade, keď uvažovaný systém umožňuje prirodzený dialóg so zmiešanou iniciatívou, kde používateľ neodpovedá iba poskytovaním informácie, a teda prevažne oznamovacími vyjadreniami, ale môže tiež klást' otázky systému a pod. Zatiaľ čo v minulosti nebolo možné nasadzovať komunikačné systémy s tak pokročilými technológiami rozpoznávania a porozumenia reči, ktoré by dovoľovali používateľovi používať prirodzený jazyk (a teda aj klást' otázky), dnes sme svedkami rapídneho zlepšenia týchto technológií, čo so sebou prináša nové výskumné a aplikačné výzvy. Jednou z nich je aj detekcia typu vyjadrenia používateľa. Rozdelenie typov vyjadrení môže byť vykonávané viac všeobecne, čím sa myslí kategorizácia podľa postojovej modálnosti prehovoru, alebo môžu byť vyjadrenia používateľa klasifikované viac detailnejšie. Vtedy hovoríme o klasifikácii na tzv. rečové, resp. dialógové akty, ktoré bližšie špecifikujú komunikačnú funkciu prehovoru používateľa.

Pre jednoduchšie komunikačné rozhrania môže byť klasifikácia prehovorov používateľov podľa postojovej modálnosti dostačujúca na identifikovanie komunikačného zámeru hovoriaceho. Komunikačné zámery hovoriaceho v súvislosti s vetným typom môžu byť rozklasifikované nasledovne:

1. Poskytnutie informácie – realizované oznamovacími vetami.
2. Získanie informácie – realizované opytovacími vetami.
3. Požiadavka na vykonanie akcie – realizované oznamovacími, rozkazovacími a želacími vetami.

Úloha automatickej klasifikácie viet podľa modálnosti pozostáva z definovania správnej sady príznakov, prípravy anotovanej databázy a z voľby vhodného klasifikačného algoritmu.

Pre voľbu správnych príznakov je potrebná analýza vlastností viet vzhľadom na postojovú modálnosť. Jednotlivé vetné typy sa líšia typom kadencie, syntaktickou stavbou a ďalším možným príznakom je aj tvar vetného prísudku. Oznamovacie vety sú typické klesajúcou kadenciou a prísudkom v indikatíve alebo kondicionáli. Opytovacie vety majú stúpajúcu kadenciu a sú pre nich typické opytovacie zámená alebo opytovacie častice. Želacie vety charakterizujú želacie častice a podmieňovací spôsob. Rozkazovacie vety majú zasa prísudok v imperatíve. [2]

Takáto skladba vlastností predikuje úspešnosť prozodických príznakov reflektujúcich kadenciu a tiež slovných príznakov, ktoré reflektujú použitie špecifických slov resp.



špecifických tvarov slovies pre úlohu automatickej klasifikácie prehovorov používateľov.

V literatúre je možné nájsť mnoho rôznych typov parametrizácií. Významná časť výskumníkov popisuje systémy spoliehajúce sa iba na parametre, ktoré sú priamo odvodené od slov v prehovore používateľa. Takýto prístup je dostatočne výkonný pre jazyky ako napríklad angličtina, kde syntaktická stavba prehovoru determinuje jeho postojovú modálnosť. Zväčša novšie práce pridávajú k slovným parametrom aj prozodické parametre, čo zvykne priniesť ďalšie zlepšenie klasifikačných systémov. Príkladom môžu byť práce [3], [4]. Je možné tiež nájsť experimenty, kde sú použité iba prozodické parametre, napr. [3], [5].

Pre natréňovanie vzorov, resp. modelov pre klasifikáciu jednotlivých vetných typov je potrebná databáza viet, ktorá pozostáva z nahrávok viet, ich prepisov a označenia ich typu. Nakoľko prozodické vlastnosti viet sú jazykovo-závislé, je potrebné pripraviť túto databázu pre konkrétny jazyk.

V našej práci sme sa rozhodli overiť výkonnosť extrémne jednoduchého klasifikátora, založeného na metóde používanej v rozpoznávaní reči – Dynamic Time Warping (DTW), pričom využívame iba jeden prozodický parameter, ktorým sú hodnoty frekvencie základného tónu hovoriaceho,  $F_0$ . Nakoľko nemáme vedomosť o existencii databázy s označenými typmi viet, rozhodli sme sa pre vytvorenie takejto databázy.

Článok je členený nasledovne. Druhá kapitola je venovaná parametrizácii použitej v úlohe klasifikovania vetných typov. V tretej kapitole je krátko popísaný použitý klasifikačný princíp na báze DTW a zdôvodnenie jeho výberu. Nasledujúca kapitola popisuje vytvorenú databázu viet. Za ňou nasleduje kapitola s výsledkami evaluačného experimentu s navrhnutým systémom.

## II. PARAMETRIZÁCIA

Melódia vety, ako kľúčová prozodická vlastnosť, je v úzkom súvisi s postojovou modálnosťou vety. Melódia vety je generovaná zmenami výšky základného tónu reči. Základný tón reči, resp. základná perióda reči je perióda hlasivkových impulzov hovoriaceho, ktoré vznikajú pri znelom buzení vokálneho traktu. Obrátenou hodnotou základnej periódy reči je tzv. frekvencia základného tónu, označovaná ako  $F_0$ . Stredná hodnota tejto frekvencie je považovaná za výšku hlasu hovoriaceho. Frekvencia základného tónu človeka je závislá od mnohých faktorov a pohybuje sa v rozmedzí od 50-400 Hz.

Priebeh frekvencie základného tónu, zostavený jej meraním na znelých úsekoch reči, naprieč celým vyjadrením používateľa je možné považovať za priebeh melódie vety. Z tohto dôvodu je vektor hodnôt  $F_0$  vhodnou základnou parametrizáciou pre systém klasifikujúci vetné typy. Viac sofistikované klasifikačné systémy pridávajú k tejto základnej parametrizácii zväčša ďalšie štatistiky vektora  $F_0$ .

Schopnosť klasifikovať vetné typy podľa priebehu melódie vychádza z predpokladu, že melódie vyjadrení používateľa v rámci jedného typu sú si veľmi podobné a zároveň pre rôzne typy odlišné.

Typy melódie môžeme rozdeliť nasledovne na [6]:

- **Konkluzívnu kadenciu.** Ide o uspokojivo končiacu melódiu vetného úseku, kde posledný vetný úsek má výrazne klesavý priebeh, teda tón poslednej slabiky je najnižší. Je typická pre oznamovacie vety.
- **Antikadenciu.** Ide o neuspokojujúco uzavretú melódiu vetného úseku, ktorá má stúpavý alebo stúpavo klesavý priebeh. Tento typ kadencie je typický pre melódiu zisťovacích opytovacích viet.
- **Semikadenciu** (polokadenciu). Je to typ melódie pred nekončiacou pauzou vo výpovedi používateľa, ktorá naznačuje, že výpoveď ešte neskončila, počúvajúcí očakáva jej pokračovanie.

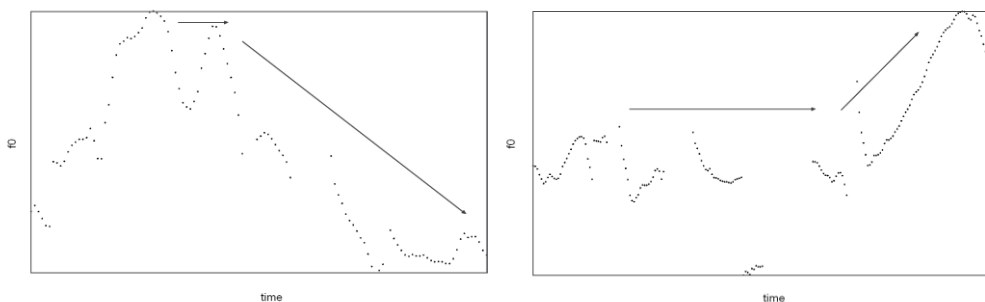


Fig. 1 Príklad konkluzívnej kadencie oznamovacej vety (vľavo) a antikadencie zisťovacej otázky (vpravo).

Aj keď niektoré typy kadencie sú typické pre viacero vetných typov, je možné v priebehu melódie badať rozdiely v priebehu, jeho strmosti a smere, ktoré by mali umožniť klasifikáciu vetných typov. Na obrázku 1 uvádzame porovnanie konkluzívnej kadencie a antikadencie.

Pre získanie vektora priebehu frekvencie základného tónu sme zvolili základný spôsob výpočtu založený na hľadaní lokálneho maxima autokorelačnej funkcie na relevantnom intervale, zodpovedajúcom fyziologicky možným hodnotám F0. Hodnoty F0 boli vypočítané pre rámce signálu s dĺžkou 30ms, pričom išlo o rozdelenie signálu na rámce bez prekryvania. Hodnoty vektora pre rámce, kde sa nachádzal rečový signál s neznelým budením boli z konečného vektora vynechané. Následne bol finálny vektor F0 normovaný, aby sa minimalizoval vplyv rečníka a tiež rozdiely pre mužských a ženských rečníkov.

### III. METÓDA DTW

Pre systém detekcie typu vety sme volili metódu DTW (Dynamic Time Warping) [7], používanú v úlohe rozpoznávania izolovaných slov, pre systémy, ktoré sú závislé na rečníkovi. Ide o algoritmus, ktorý umožňuje meranie podobnosti medzi dvoma sekvenciami príznakov, v prípade, že tieto popisujú signály, ktoré variujú v čase alebo rýchlosti. Algoritmus DTW je teda schopný vypočítať vzdialenosť dvoch signálov, ktoré sa líšia v svojej dĺžke, a teda aj v počte príznakov. V úlohe rozpoznávania reči je táto premenlivosť v dĺžke odrazom rozdielnej dĺžky vyslovených slov. V našej úlohe, premenlivosť dĺžky reprezentuje premenlivú dĺžku vety a teda aj príznakového vektora.

Princípom metódy DTW je, že časové rozdiely medzi porovnávanými sekvenciami sú vyrovnávané nelineárnou transformáciou časovej osi, kde toto vyrovnanie sa riadi podľa vzdialeností medzi jednotlivými hodnotami resp. vektormi.

Metóda DTW, ako aj celý systém pre detekciu typu prehovoru, bola implementovaná v Matlabe. Najprv bola zvolená skupinu vzorov (patterns), ktoré reprezentujú jednotlivé typy viet – oznamovacie, opytovacie - zisťovacie, opytovacie - doplňovacie a rozkazovacie. Pre každú skupinu sme zvolili osem vzorov, s ktorými algoritmom DTW porovnávajú testované vety. Každý testovaný súbor je parametrizovaný a je vypočítaná jeho vzdialenosť ku každému zo vzorov. Následne sú tieto vzdialenosti vyhodnotené a nahrávka sa označí tým typom vety, ku ktorej skupine vzorov má najmenšiu vzdialenosť.

### IV. DATABÁZA VIET

Výhodou popísaného prístupu k detekcii typu vety je, že nepotrebuje pre vytvorenie vzorov rozsiahlu databázu. Napriek tomu, pre potreby výskumu a prípadného natrénovania modelov pre metódu na báze skrytých Markovových modelov (HMM) a tiež pre účely otestovania zvolenej metódy sme pristúpili k vytvoreniu malej databázy s označením typu viet. Databáza pozostáva z časti, ktoré tvoria vlastné nahrávky a časti, ktorá vznikla rozsegmentovaním a oantovaním časti relácie „Pod lampou“. Celkový počet rečníkov v databáze je 40. Počet nahrávok pre jednotlivé časti databázy sa nachádza v tabuľke 1.

Tabuľka 1  
 Počet nahrávok databázy typov viet

	Oznamovacie vety	Rozkazovacie vety	Opytovacie vety
Počet nahrávok z relácie „Pod lampou“	147	0	227

Počet vlastných nahrávok	103	99	173
Celkový počet nahrávok	250	99	400

## V. VÝSLEDKY

Pre úlohu detekcie typu vety bolo potrebné najprv zvoliť vhodnú sadu vzorov. Pre každý typ vety sme vybrali 4 rôzne vzory, ktoré sa líšili dĺžkou, nakoľko je zrejme, že metóda DTW bude efektívna iba do istého rozdielu dĺžok. V prípade opytovacích viet sme zvolili až 6 vzorov, nakoľko sme potrebovali zastúpiť tak doplňovacie ako aj zisťovacie otázky s rôznymi dĺžkami trvania.

Úspešnosť rozpoznania jednotlivých typov viet a celková úspešnosť je zaznamenaná v Tab. 2.

Tabuľka 2  
 Výsledky evaluačného experimentu 1

	Oznamovacie vety	Rozkazovacie vety	Opytovacie vety
Spôľahlivosť (Accuracy) [%]	64,37	52,54	67,87
Celková spoľahlivosť [%]	<b>61,59</b>		

Ako je zrejme z tabuľky č. 2, úspešnosť detekcie rozkazovacích viet je nízka. Analýzou databázy sme dospeli k dvom záverom. Nakoľko časť databázy, ktorá obsahuje nahrávky z relácie „Pod lampou“ neobsahovala rozkazovacie vety, databáza obsahuje iba imitované rozkazovacie vety, ktoré boli nahrané študentmi. Po ich vypočutí je možné konštatovať, že ich melódia skutočne nezodpovedá reálnym rozkazovacím vetám. V niektorých prípadoch ide skôr o vety oznamovacie, iba s väčším dôrazom v prvej časti vety. Druhým faktorom je, že použitie iba prozodického parametra v podobe vektora F0 je nedostatočné pre rozlíšenie oznamovacích a rozkazovacích viet, nakoľko sú obe typické konkluzívnou kadenciou. Preto sme pristúpili k druhému experimentu, kde sme rozkazovacie vety z experimentu vylúčili (viď. Tab.3)

Tabuľka 3  
 Výsledky evaluačného experimentu 2 – bez rozkazovacích viet

	Oznamovacie vety	Opytovacie vety
Spôľahlivosť (Accuracy) [%]	87,67	67,87
Celková spoľahlivosť [%]	77,77	

Vylúčenie rozkazovacích viet zlepšilo detekciu oznamovacích viet, čo bol očakávaný výsledok. Celkovo je možné zhodnotiť, že vzhľadom na skalárny typ parametrizácie (na segment) a jednoduchý klasifikačný prístup je celková spoľahlivosť systému dobrá a môže byť výrazne zlepšená pridaním slovných parametrov.

## VI. ZÁVER

Predložený článok popisuje problém detekcie vetných typov v prehovoroch používateľa z hľadiska postojovej modálnosti a tiež systém, ktorý bol navrhnutý pre túto úlohu. Prezentovaný systém je založený na meraní vzdialeností metódou DTW medzi vektormi priebehov základného tónu jednotlivých prehovorov používateľa a sadou vzorov, ktoré reprezentujú jednotlivé vetné typy. Ide o jednoduchý systém, ktorý je výpočtovo nenáročný a nespolieha sa na prepis prehovoru systémom automatického rozpoznávania reči. Môže teda slúžiť ako prvotná informácia o vetnom type, ktorá môže byť korigovaná na základe slovného parametra získaného z prepisu prehovoru.

Prezentovaný systém môže byť nasadený v rečových komunikačných systémoch, kde môže pomôcť v rozhodovaní dialógového manažéra o vhodnej reakcii na používateľský vstup.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných

technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja (50%) a projektu Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií (ITMS-26220220182) 50% na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### REFERENCIE

- [1] Páleš, E. SAPFO, Parafrázovač slovenčiny. Počítačový nástroj na modelovanie v jazykovede. Bratislava: Veda, 1994. 308 s. ISBN 80-224-0109-9.
- [2] Páleš, E., Sémantické roly slovenských slovies, Jazykovedný časopis, vol. 41, č. 1, 1990, str. 30 – 46.
- [3] Vu Q. Minh, L. Besacier, E. Castelli, Automatic question detection: prosodic-lexical features and crosslingual experiments. In Interspeech 2007-Eurospeech, 2007
- [4] Shriberg, E., et al., “Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech?” Language and Speech 41, pp. 439-487, 1998.
- [5] S. O. Olatunji, L. Cheded, W. G. Al-Khatib, O. Khan, Identification of Question and Non-Question Segments in Arabic Monologues Using Prosodic Features: Novel Type-2 Fuzzy Logic and Sensitivity-Based Linear Learning Approaches. In: Journal of Intelligent Learning Systems and Applications, 2013, 5, 165-175
- [6] UHLÁR, Vlado: O vetej melódii v slovenčine. In Slovenská reč. 1958, roč. 23, č.6.
- [7] R. Bellman and R. Kalaba, “On adaptive control processes”, Automatic Control, IRE Transactions on , vol. 4, no. 2, pp. 1-9,1959.

# Solving Economical Problems by Vector Optimization Methods

<sup>1</sup>Ján ČABALA, <sup>2</sup>Anna JADLOVSKÁ

Department of Cybernetics and Artificial Intelligence, Faculty of Electrical  
Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>jan.cabala@tuke.sk, <sup>2</sup>anna.jadlovska@tuke.sk

**Abstract** — This paper deals with solving economical problems by methods of vector optimization. Description and solving vector optimization tasks using chosen methods is the main issue. Paper also contents example of economical vector optimization problem, which can be specified as problem of choosing the ideal investment strategy. Application in MATLAB was created for intuitive and simple solving vector optimization tasks.

**Keywords** — economical problems, linear programming, quadratic programming, vector optimization

## I. INTRODUCTION

Optimization is defined as choosing the best option from the set of possible options. Optimization tasks and problems take place in many different sectors. The main goal in economical sector is to maximize profit or minimize costs. It can be used also in solving traffic problems, specifically in minimizing time and costs of transportation. Also consumers might use optimization methods to satisfy their needs as much as possible with minimum costs. Publications dealing with optimization and vector optimization task are [1] and [4]. Main asset of this work, based on [8], is creating MATLAB application for simple and intuitive solving vector optimization tasks, because there wasn't any complex application able to solve these tasks. Application provides users with user interface for solving vector optimization problems with two objective functions. Using of application is illustrated by solving problem of defining the optimal investment strategy using vector optimization methods.

## II. BRIEF OVERVIEW OF OPTIMIZATION METHODS

Optimization process is disposing with apparatus to choose the best possible solution for particular task. In order to make final decision responsibly and competently, it is necessary to:

1. create mathematical model which precisely describes the situation. Model represents mathematical description of real system. It has to contain quantifiable parameters, which are used to evaluate the success rate of optimization process (profit, costs etc.). Model can also contents some constrains (f.e. maximal amount of invested money);
2. find the solution of particular optimization problem using proper algorithm, because there is no universal algorithm, which can be used to solve all optimization tasks (tasks solved by methods of linear and quadratic programming are described in this paper);
3. verify and analyze found solution (if it is the real solution of optimization task). It is also necessary to interpret the solution correctly.[1]

General optimization problem can be defined as minimization (maximization) of objective function

$$f = \{x_1, x_2, \dots, x_n\} \quad (1)$$

with respecting all constrains

$$\begin{aligned} g_i &= \{x_1, x_2, \dots, x_n\}, & \text{for } i = 1, 2, \dots, n \\ x_j &> 0, & \text{for } j = 1, 2, \dots, n. \end{aligned} \quad (2)$$

Methods of mathematical programming are used to solve optimization tasks. Depending on the type of objective function, these methods can be divided into:

- linear programming methods,
- non linear programming methods,
- integer programming methods
- parameter programming methods,
- stochastic programming methods.

In following parts of the paper there will be a more detail description of linear and quadratic programming methods used to solve optimization tasks. Values of these solutions will be subsequently used in vector optimization process.

#### A. Linear programming task description

Linear programming is part of mathematical programming dealing with finding optimal solutions of optimization tasks.

Most of optimization tasks consists of two parts:

- main goal – defined by objective function,
- constrains .

Both of these parts are described by linear functions with multiple variables. Generally the linear programming problem can be described with following mathematical model:

$$z = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \text{opt}(\min, \max) \quad (3)$$

with respecting constrains

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &\geq \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &\geq \leq b_2, \\ &\dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &\geq \leq b_n. \end{aligned} \quad (4)$$

Other way to define the task is to use the matrix form:

$$\begin{aligned} z &= c^T x \rightarrow \text{opt}(\min, \max) \\ Ax &\geq \leq b, \end{aligned} \quad (5)$$

in which  $c$  is vertical vector of objective function coefficients,  $x$  is vertical vector of task variables,  $A$  is matrix describing left sides of constrains  $b$  is vertical vector describing right sides of constrains.[1]

#### B. Quadratic programming task description

Goal of quadratic programming task is to optimize quadratic objective function with linear constrains. To describe general form of quadratic programming task, there is a need to mention some terms inevitable for closer definition of this form.

There is a symmetrical and positively semidefinite matrix  $C \in \mathbb{R}^n$  and there is  $d, a_i \in \mathbb{R}^n, b_i \in \mathbb{R}, i = 1, 2, \dots, m$ . Let's label:

$$\begin{aligned} P &= \{x \in \mathbb{R}^n, x_1, \dots, x_n \geq 0, s \in \{0, \dots, n\}\} \\ X &= \{x \in P \langle a_i, x \rangle \leq b_i, i = 1, \dots, k; \langle a_i, x \rangle = b_i, i = k + 1, \dots, m\} \end{aligned} \quad (6)$$

After this definition we can define general form of quadratic programming task as

$$f(x) = \frac{1}{2} \langle Cx, x \rangle + \langle d, x \rangle \rightarrow \min, x \in X. \quad (7)$$

We will label  $A = [a_1, \dots, a_m]^T$ , where  $a_1, \dots, a_n \in \mathbb{R}^n$  are rows of matrix  $A$  a  $b = [b_1, \dots, b_m]^T$ . With closer knowledge it is possible to define Lagrange function of quadratic programming task:

$$L = \frac{1}{2} \langle Cx, x \rangle + \langle A^T y + d, x \rangle - \langle y, b \rangle, \quad (8)$$

and so

$$L'_x(x, y) = Cx + A^T y + d. \quad (9)$$

It is important to mention, that  $y \in Q = \{y = (y_1, \dots, y_m), y_1, \dots, y_k \geq 0\}$ .

Solution of quadratic programming task is  $x^*$  just when there is  $y^* \in Q$  respecting so called Kuhn – Tucker conditions:

$$\begin{aligned} \langle Cx^* + A^T y^* + d, x - x^* \rangle &\geq 0, \forall x \in P \\ y_i^* (Ax^* - b)_i &= 0, i = 1, \dots, m. \end{aligned} \quad [3] \quad (10)$$

### III. VECTOR OPTIMIZATION DESCRIPTION AND METHODS

Vector optimization is dealing with ways of optimizing the problem with multiple goals. It is used when it is necessary to accept more than one factor in finding the ideal solution of optimization task. This type of optimization was created for solving planning and organizing problems in manufacturing process. Nowadays it is used in many different areas (f.e. in dynamic management systems).

Vector optimization task is defined with relation to controlled system. It is described with  $n$ -dimensional vector  $(x_1, x_2, \dots, x_n)$ , with  $x \in \{X\}$  and is evaluated with  $k$ - dimensional vector functional  $J(x) = (J_1(x), J_2(x), \dots, J_k(x))$ , which elements are functions of vector  $x$ . Solution of this task is  $x^* \in \{X\}$ , which is optimal solution of functionals  $J_1(x), J_2(x), \dots, J_k(x)$  according to chosen type of vector optimization.

If we connect this definition of vector optimization to complex process  $P$  created from partial processes  $P_i, i = 1, 2, \dots, k$  (every partial process is defined by equations and constrains and is controlled by its criterion  $J_i$ ), we can claim that  $x^*$  is optimizing process  $P$  with regard to function  $R(x)$  with vector quality pointer  $J(x)$ , if is valid

$$R(x^*) = \inf R(x), \text{ with } x^* \in \{X\}. \quad (11)$$

From the 70's, when first vector optimization tasks were solved, numbers of methods dealing with this issue were created. We can divide them into two main groups:

- methods defining the set of non improving elements ,
- compromising methods.

#### A. Methods defining the set of non improving elements

There is no hierarchy of criteria in this group of methods, every criterion is equally important. Task of vector optimization can be defined as trying to minimize the vector  $J(x) - J(x_\alpha)$ , where  $x_\alpha$  is optimal values vector of variables  $x$  according to defined criteria.

##### 1) Quadratic norm

The most common criterion used in this norm is minimal sum of quadratic variance of objective functions  $J_\alpha(x)$  for random  $x \in \{X\}$  from objective function  $J_\alpha(x_\alpha)$  for vector of ideal values of chosen criteria  $x_\alpha = (x_{1\alpha}, x_{2\alpha}, \dots, x_{n\alpha}), \alpha = 1, 2, \dots, k$ ;

$$R(x) = \sum_{\alpha=1}^k (J_\alpha(x) - J_\alpha(x_\alpha))^2$$

$$\text{opt } R(x) = \min_{x \in \{X\}} \sum_{\alpha=1}^k (J_\alpha(x) - J_\alpha(x_\alpha))^2 \quad (12)$$

Values of this norm are usually divided by optimal value. Reason for this is non-dimensional solution value. Calculation formula then looks like:

$$R(x) = \sum_{\alpha=1}^k \frac{(J_\alpha(x) - J_\alpha(x_\alpha))^2}{J_\alpha(x_\alpha)^2}$$

$$\text{opt } R(x) = \min_{x \in \{X\}} \sum_{\alpha=1}^k \frac{(J_\alpha(x) - J_\alpha(x_\alpha))^2}{J_\alpha(x_\alpha)^2}. \quad (13)$$

##### 2) Linear norm

It represents minimal sum of linear variance of objective functions  $(J_\alpha(x))$  from optimal values of objective functions  $J_\alpha(x_\alpha)$ . Optimal value of functions  $R(x)$  can be calculated using formulas:

$$R(x) = \left| \sum_{\alpha=1}^k (J_\alpha(x) - J_\alpha(x_\alpha)) \right|$$

$$\text{opt } R(x) = \min_{x \in \{X\}} \left( \sum_{\alpha=1}^k (J_\alpha(x) - J_\alpha(x_\alpha)) \right). \quad (14)$$

##### 3) Generalized norm

Objective functions of this norm is

$$R_L(x) = \sum_{\alpha=1}^k \left\{ (J_\alpha(x) - J_\alpha(x_\alpha))^L \right\}^{\frac{1}{L}}; L \geq 1. \quad (15)$$

For  $L = 1$  this functions corresponds to linear norm, for  $L = 2$  it corresponds quadratic norm and for  $L = \infty$  it is

$$R_{\infty}(x) = \max_{\beta} \{J_{\alpha}(x) - J_{\alpha}(x_{\alpha}); \alpha = 1, 2, \dots, n\}. \quad (16)$$

All these norms can be refined by multiplying objective function elements with appropriate coefficients. Searched point  $x^* \in \{X\}$  is called non improving in space  $\{X\}$  regarding to functional  $J(x)$ , if there is no point  $\tilde{x}$  in this space, for which is valid  $J_{\alpha}(\tilde{x}) \leq J_{\alpha}(x^*), \alpha = 1, 2, \dots, k$ .

### B. Compromising methods

These methods are based on defining more strict constrains or adding another constrains to objective functions. Compromise means finding optimal solution by minimizing value of formula  $\beta_1 J_1(x) + \beta_2 J_2(x) + \dots + \beta_k J_k(x)$ , where  $\beta_1$  to  $\beta_k$  are carefully chosen importance coefficients. Their values are recommended to be chosen as follows:  $\beta_1 = 1/J_{10}, \beta_2 = 1/J_{20}, \beta_k = 1/J_{k0}$ , where  $J_{k0}$  are values of objective functions counted in optimization only via selected criterion.

These methods are used when it is possible to define importance of every criterion before starting optimization process, or when additional information about criterion importance is found out during optimization process. Main idea is in defining importance of every scalar criterion. This can affect the result of vector optimization (importance of scalar criterion will be labeled as  $\lambda$ ). Most common methods from this group are optimization of weighted sum of scalar criteria and weighted sum of variance vector.

#### 1) Optimization of weighted sum of scalar criteria

Function for this norm is defined by formula

$$\text{opt } R(x) = \text{opt} \sum_{\alpha=1}^k \lambda_{\alpha} J_{\alpha}(x). \quad (17)$$

Optimal solution is represented by maximum or minimum of weighted sum of scalar criteria, depending on character of problem being solved.

#### 2) Weighted sum of variance vector

In this norm function  $R(x)$  is represented by formula

$$R(x) = \sum_{\alpha=1}^k \lambda_{\alpha} * (J_{\alpha}(x) - J_{\alpha}(x_{\alpha})). \quad (18)$$

Optimal solution is always represented by minimum value of function  $R(x)$ , because this norm is using variation from ideal values.

$$\text{opt } R(x) = \min_{x \in \{X\}} \left( \sum_{\alpha=1}^k \lambda_{\alpha} * (J_{\alpha} - J_{\alpha}(x_{\alpha})) \right). \quad (19)$$

It is also possible to use quadratic variant of this formula (weighted quadratic norm).[4]

## IV. CHOOSING OPTIMAL INVESTMENT STRATEGY BY USING VECTOR OPTIMIZATION METHODS

### A. Defining the problem

Investor was given some capital from the lender and he is deciding about the best investment of the funds. He can divide whole capital into three commodities. Investor has economical knowledge of Markowitz model (closely described in [5] and [6]) and his extension for safe investment. At the basis of these knowledge investor defines main conditions of his strategy:

Investor expects minimum profit of 10%. He assumes 10% profitability from investment to 1<sup>st</sup> commodity, 5% profitability from investment to 2<sup>nd</sup> commodity a 16% profitability from investment to 3<sup>rd</sup> commodity. Investor also knows variance  $\sigma_i^2$  of  $i^{\text{th}}$  commodity and covariance coefficients  $\sigma_{ij}$   $i^{\text{th}}$  and  $j^{\text{th}}$  commodity. Values of these coefficients are:  $\sigma_1^2 = 0,2; \sigma_2^2 = 0,03; \sigma_3^2 = 0,18; \sigma_{12} = 0,05; \sigma_{13} = 0,02; \sigma_{23} = 0,03$ . Investor's goal is to minimize the riskiness of his investment. This can be reached by minimizing the variance of commodity portfolio. It is defined by formula  $Var_p = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + x_3^2 \sigma_3^2 + 2x_1 x_2 \sigma_{12} + 2x_1 x_3 \sigma_{13} + 2x_2 x_3 \sigma_{23}$ . [6]

Lender expects refunding of investment by monthly payments. He monthly claims 8% from capital invested to 1<sup>st</sup> commodity, 10% from capital invested to 2<sup>nd</sup> commodity and 7% from capital invested to 3<sup>rd</sup> commodity. Investor also doesn't want to invest more than 60% of given capital to one commodity.



What is the ideal investment strategy for investor in order to minimize the riskiness of investment and also minimize monthly payment to the lender?

*B. Solution*

This task is vector optimization task. Solution is to find ideal values of vector  $x$  by optimizing objective functions

$$\begin{aligned} 0,2 x_1^2 + 0,08 x_2^2 + 0,18 x_3^2 + 0,1x_1x_2 + 0,04x_1x_3 + 0,06x_2x_3 &\rightarrow \min \\ 0,08x_1 + 0,1x_2 + 0,07x_3 &\rightarrow \min \end{aligned} \quad (20)$$

with respecting constrains

$$\begin{aligned} 0,1x_1 + 0,05x_2 + 0,16x_3 &\geq 0,1 ; \\ x_1 + x_2 + x_3 &= 1; \\ 0 \leq x_1, x_2, x_3 &\leq 0,6. \end{aligned} \quad (21)$$

First step is find optimal solutions of objective functions separately, so just for

$$0,08x_1 + 0,1x_2 + 0,07x_3 \rightarrow \min$$

and

$$0,2 x_1^2 + 0,08 x_2^2 + 0,18 x_3^2 + 0,1x_1x_2 + 0,04x_1x_3 + 0,06x_2x_3 \rightarrow \min.$$

Result of this optimization is displayed in Table 1. From the results we can claim, that in optimizing for monthly payment investor will invest most of his funds to 3<sup>rd</sup> commodity and he won't invest any of funds to 2<sup>nd</sup> one. On the other hand, in minimizing riskiness ratio, most of funds will be invested to 2<sup>nd</sup> commodity, and the smallest part will be invested to 1<sup>st</sup> one.

Table 1 Optimizing for separate objective functions

	$x_1$	$x_2$	$x_3$
Optimization for monthly payment	40%	0%	60%
Optimization for riskiness ratio	19,447%	43,938%	36,615%

Another step is calculation of ideal values for both objective functions. This calculation is done by inserting ideal values counted before into objective functions. After this calculation it is clear that minimal monthly payment for lender is 7,4 % from invested funds and minimal portfolio variation is 6,82% ( it means that standard deviation is 26,11 % ) . After this step the calculation of vector optimization using quadratic norm is performed. Goal of this step is to minimize variation from ideal values of both objective functions:

$$\text{opt } R(x) = R(x^*) = \min \sum_{b=1}^m \frac{(J_b(x) - J_b(x_b))^2}{J_b(x_b)^2}. \quad (22)$$

After inserting coefficients form objective functions this equation is created:

$$\begin{aligned} R(x) = &\left( \frac{0,08x_1 + 0,1x_2 + 0,07x_3}{740} - 1 \right)^2 + \\ &+ \left( \frac{0,2 x_1^2 + 0,08 x_2^2 + 0,18 x_3^2 + 0,1x_1x_2 + 0,04x_1x_3 + 0,06x_2x_3}{1363700} - 1 \right)^2. \end{aligned} \quad (23)$$

After the calculation minimal value of variance is found as well as result of vector optimization (ideal vector  $x$  using quadratic norm).

Table 2 Optimal percentage distribution of given capital between commodities

	$x_1$	$x_2$	$x_3$
Optimal values	22,892%	36,573%	40,535%

In this distribution monthly payment represents 8,33 % of lent funds and portfolio variance is 7,17% (standard deviation is 26,78%).

Table 3 displays comparison of vector optimization results using various vector optimization methods, specifically linear norm, 4<sup>th</sup> power norm, and weighted quadratic norm for different importance of objective functions were used (1<sup>st</sup> value of importance is for optimization of monthly payment, 2<sup>nd</sup> one is for optimization of riskiness ratio).

Table 3 Ideal distribution values using other norms of vector optimization

	$x_1$	$x_2$	$x_3$
optimization - monthly payment	40%	0%	60%
optimization - riskiness ratio	19,447%	43,938%	36,615%
VO - quadratic norm	22,892%	36,573%	40,535%
VO - linear norm	19,447%	43,938%	36,615%
VO - 4 <sup>th</sup> power norm	22,373%	37,681%	39,945%
VO - weighted quadratic norm (0,4 ; 0,6)	19,447%	43,938%	36,615%
VO - weighted quadratic norm (0,6 ; 0,4)	26,206%	29,489%	44,305%
VO - weighted quadratic norm (0,7 ; 0,3)	37,315%	5,740%	56,945%

Results are also displayed on the graph in Fig. 1.

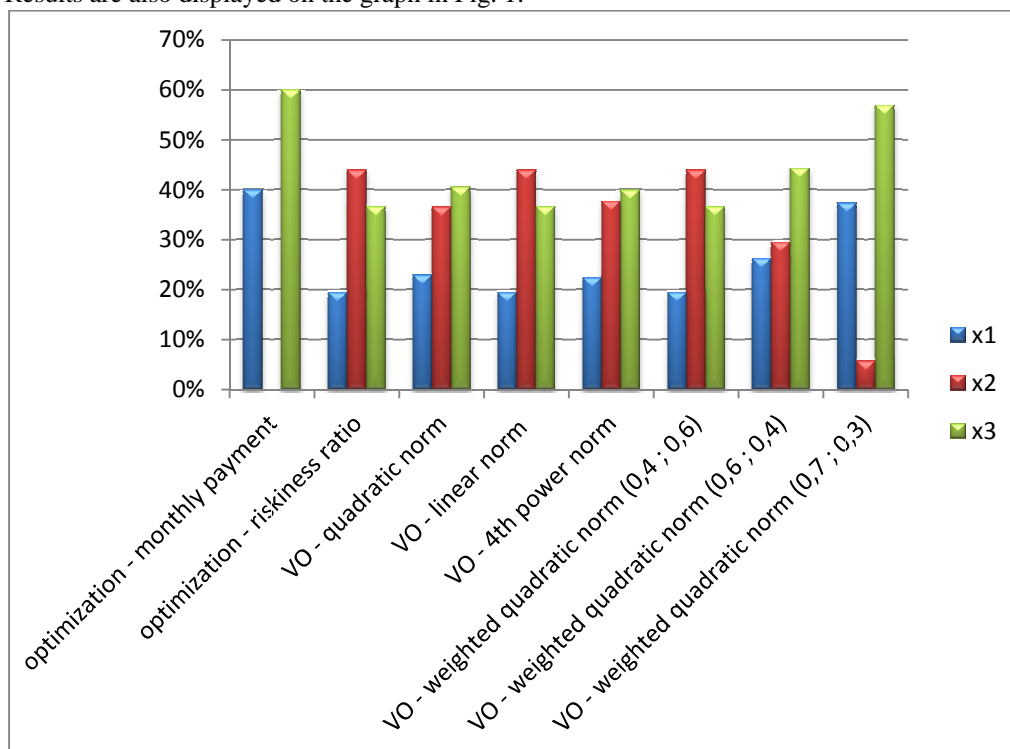


Fig. 1 Percentage of investment distribution into commodities using various methods of vector optimization

Application in MATLAB was created for solving vector optimization tasks. It can be used for solving vector optimization tasks for 2 objective functions with linear or quadratic form. More specific description of this application can be found in [8]. Applications interface is shown in Fig. 2.

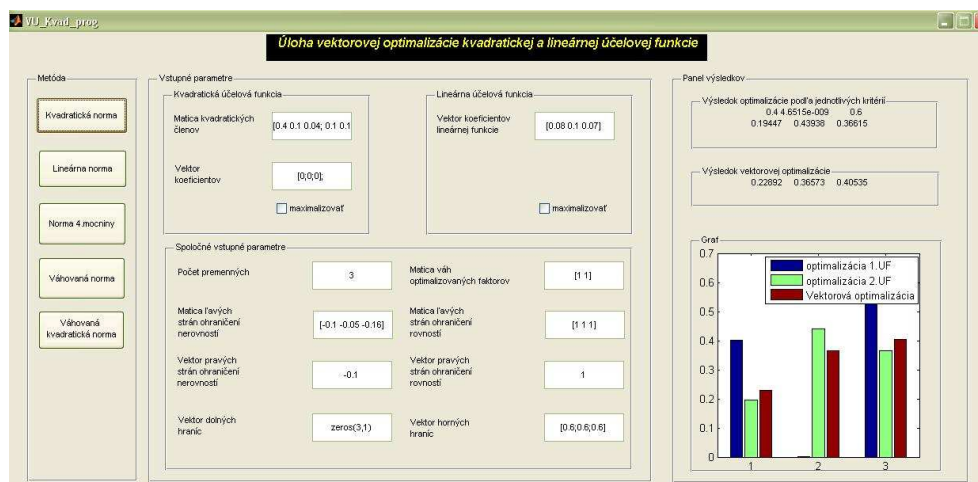


Fig. 2 Application for solving vector optimization problems

## V. CONCLUSION

Vector optimization represents simple and practical way to solve optimization task in case that there is more than one optimized factor. Many organizations face vector optimization problems, mainly when they are trying to reach optimal process functionality in organization. This paper contains various methods of vector optimization. In conclusion we can claim that results of vector optimization mainly depend on chosen method of optimization. In most cases the result of vector optimization represents some kind of compromise between partial optimizations for separate objective functions.

## ACKNOWLEDGEMENT

This contribution is the result of the Vega project implementation – Dynamic Hybrid Architectures of the Multi-agent Network Control Systems(No.1/0286/11), supported by the Scientific Grant Agency of Slovak Republic – 40%, and of the KEGA project implementation – CyberLabTrainSystem – Demonstrator and Trainer of Information-Control Systems (No. 021TUKE-4/2012) – 30%, and of the Research and Development Operational Program for project: University Science Park Technicom for innovative applications with knowledge technology support, (ITMS code 26220220182), co-financed by the ERDF – 30%.

## REFERENCES

- [1] ROSINOVÁ, Danica – DÚBRAVSKÁ Mária: Optimization. Vol.1, Bratislava: STU, 2008. 190 pages. - ISBN: 8022727952.
- [2] BEREŽNÝ, Štefan – KRAVECOVÁ, Daniela: Linear programming. Vol.1. Košice: KMTI FEI TU, 2012. 108 s.- ISBN 978-80-553-0910-1.4
- [3] JIŘÍČKOVÁ, Viera: Chosen methods of quadratic programming: Diploma thesis. Brno: ÚMaS PF. 2011. 78 pages.
- [4] HRUBINA, Kamil - JADLOVSKÁ Anna – HREHOVÁ Stella: Algorithms of optimization methods with using programming systems. Vol.1 . Prešov – Košice, 2005.393 pages. ISBN 80-88941-31-8.
- [5] PEJCHAL, Lukáš: Defining optimal weights of funds portfolio: Diploma thesis. Brno: ÚMaS PF. 2011. 73 pages.
- [6] INUIGUCHI, Masahiro – TANINO Tetsuzo: Portfolio selection under independent possibilistic information. Fuzzy Sets and Systems, 2000, pages 83-92. Available on internet:< <http://www.geocities.ws/smhurtado/2000-115-1-83-92.pdf>>
- [7] SKÁCEL, Ján : Riskiness reduction possibilities of portfolio investments: Diploma thesis. Bratislava: KFaE FM UK. 2005. 58 s. Available on internet: < <http://diplomovka.sme.sk/zdroj/2737.pdf> >.
- [8] ČABALA, Ján: Solving economical problems by methods of multiple-criteria decision-making and vector optimization: Diploma thesis. Košice: KKUI FEI TUKE. 2014. 89 pages.

# Simulácia teploty výkonových prvkov v elektronických štruktúrach s využitím softvéru HyperLynx® Thermal

*Pavol CABÚK, Tomáš GIRAŠEK, Slavomír KARDOS*

<sup>1</sup> Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

pavol.cabuk@tuke.sk, slavomir.kardos@tuke.sk

**Abstrakt** — Súčasný trend miniaturizácie elektronických zariadení prináša množstvo nových problémov. Jedným z nich sú aj zvyšujúce sa nároky na tepelný manažment, keďže zvýšenie stupňa integrácie má za následok nárast výkonovej hustoty a teda aj pracovnej teploty komponentov, čo prináša problémy so spoľahlivosťou a životnosťou. Najefektívnejším riešením týchto problémov je ich minimalizácia už vo fáze návrhu s využitím súčastí návrhových systémov umožňujúcich tepelnú analýzu navrhovaného zariadenia. Príspevok sa venuje porovnaniu výsledkov simulácií softvérovým nástrojom HyperLynx® Thermal s reálnymi vzorkami. Výsledky poukazujú na praktické obmedzenia použitého softvérového nástroja s posúdením korelácií simulácií s meraniami termokamerou FLIR E30.

**Kľúčové slová**— CAD, HyperLynx® Thermal, optimalizácia návrhu, teplotné simulácie

## I. ÚVOD

Funkcionalita teplotných simulácií nadobúda na dôležitosti spolu so stále sa zosilňujúcim trendom miniaturizácie, keď aj pri pôvodne bezproblémových konštrukciách dochádza k podstatnému nárastu výkonovej hustoty, a tým aj k problémom s nárastom pracovnej teploty. Pracovná teplota komponentov je jedným z hlavných faktorov vplývajúcich na spoľahlivosť zariadenia. Analýza rozloženia teploty na DPS už v priebehu návrhu prináša možnosť včasnej optimalizácie návrhu. Vďaka možnosti simulácie a vizualizácie rozloženia teploty na DPS je možné predikovať problémové oblasti a vhodným konštrukčným zásahom zlepšiť odvod tepla, a teda aj znížiť pracovnú teplotu komponentov a tepelné namáhanie substrátu. Vďaka tomu je možné optimalizovať pracovnú teplotu komponentov už vo fáze návrhu, čo prináša zásadné zvýšenie spoľahlivosti elektronických zariadení pri minimalizovaných nákladoch potrebných na modifikáciu prototypu. [1]

## II. SIMULÁCIA TEPLOTNÝCH POLÍ

Pre simuláciu teplotných polí v elektronických zariadeniach je dostupných viacero profesionálnych softvérových nástrojov. Od jednoduchých, poskytujúcich len základnú analýzu v statickom režime, cez komplexnejšie, umožňujúce simuláciu prechodových a dynamických dejov, až po programy pracujúce s podrobnými fyzikálnymi modelmi, umožňujúce vernú simuláciu vrátane vzájomného ovplyvňovania komponentov konvexiou a radiáciou.

Výhodou jednoduchších je rýchla orientačná simulácia teplotných pomerov umožňujúca odhalenie rizikových miest s prílišným nahustením výkonových prvkov. Programy tohto typu je možné jednoducho prepojiť s najpoužívanejšími CAD systémami pre návrh DPS. V mnohých prípadoch sú priamo súčasťou týchto systémov, aj keď sa zväčša jedná o osobite platené moduly. Spravidla obsahujú dostatočne presné modely používaných komponentov. Jedným z hlavných obmedzení je práve jednoduchosť simulácií, ktoré neumožňujú podchytiť komplexnejšie deje ako napríklad odraz IČ žiarenia od okolitých plôch. Ich význam oslabuje aj fakt, že skúsenejší konštruktér vie do značnej miery odhadnúť kritické miesta navrhovaného zariadenia aj bez potreby simulačného nástroja.

Komplexnejšie nástroje sa uplatňujú pri simuláciách zariadení, ktoré sú určené pre prácu v zložitejších podmienkach alebo v režime, kde dochádza ku krátkodobému preťažovaniu

zariadenia. V takýchto prípadoch je nutné skúmať teplotné polia v dynamickom režime a uvažovať aj tepelné kapacity, zmeny prúdenia atmosféry následkom ohrevu a podobne. Tento typ simulačných programov je spravidla distribuovaný ako samostatný program s možnosťou importu z iných CAD systémov.

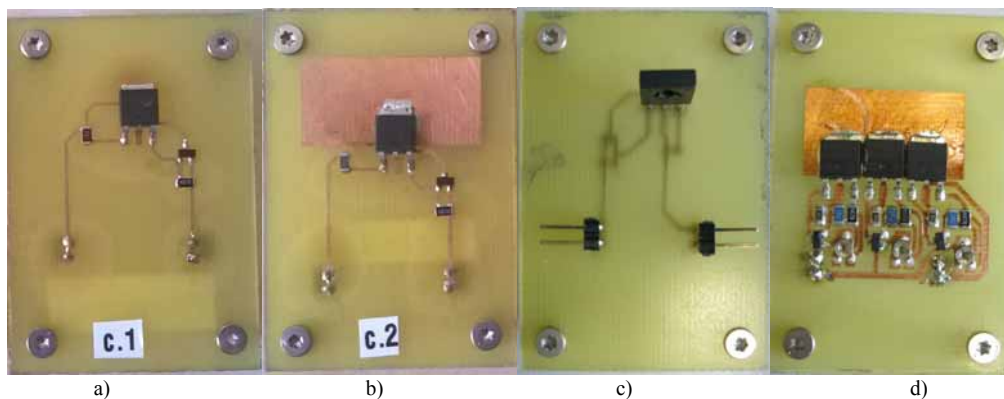
Poslednú, najvyššiu triedu predstavujú nástroje schopné simulácie na základe podrobných fyzikálnych modelov. Spravidla pracujú na základe metódy konečných prvkov s detailnými fyzikálnymi modelmi komponentov. Poskytujú obraz dobre korešpondujúci so skutočnosťou za predpokladu správneho zostavenia modelov..

### III. SOFTVÉROVÝ NÁSTROJ HYPERLYNX® THERMAL

Softvér HyperLynx® Thermal od firmy MentorGraphic® umožňuje tepelné analýzy DPS vo fáze CAD návrhu. Umožňuje import z najpoužívanejších návrhových systémov. Knižnice obsahujú štandardne používané materiály (FR4 a pod.) a databázu materiálov pre zostavenie substrátu podľa špecifikácie návrhára. Je možné pracovať s DPS ľubovoľného tvaru, pričom softvér pri výpočtoch zohľadňuje umiestnenie dosky. Umožňuje simulovať jednotlivé voľne uložené DPS, DPS na podložke alebo v definovanom mieste vyššieho konštrukčného celku. Softvér taktiež zohľadňuje spôsob ukotvenia dosky pomocou vodivých alebo nevodivých podložiek a skrutiek v DPS alebo v chladičoch. Tým sú zohľadňované konštrukciou dané tepelné mosty a spôsob obtekania vzduchom vrátane vplyvu gravitácie, tlaku a núteného prúdenia vzduchu. Softvér umožňuje prácu s ustáleným prúdením vzduchu nastaviteľným vo všetkých 3 osiach s celkovou presnosťou modelovania vplyvu konvekcie a radiácie  $\pm 10\%$ . Táto presnosť bola experimentálne overená pri vývoji softvéru v aerodynamickom tuneli za pomoci infračervenej kamery [2].

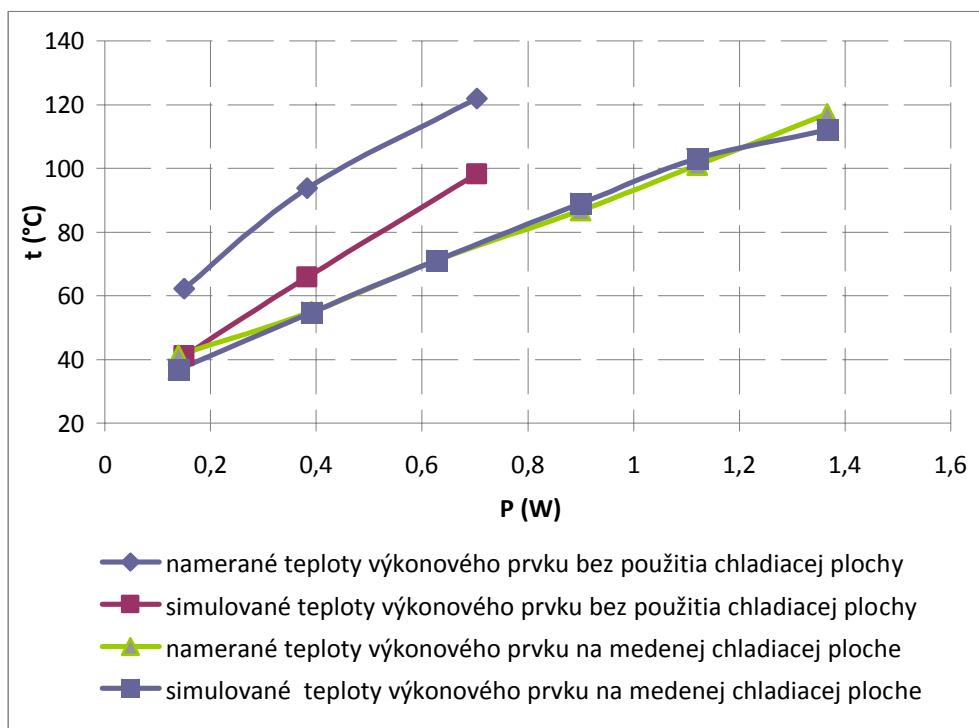
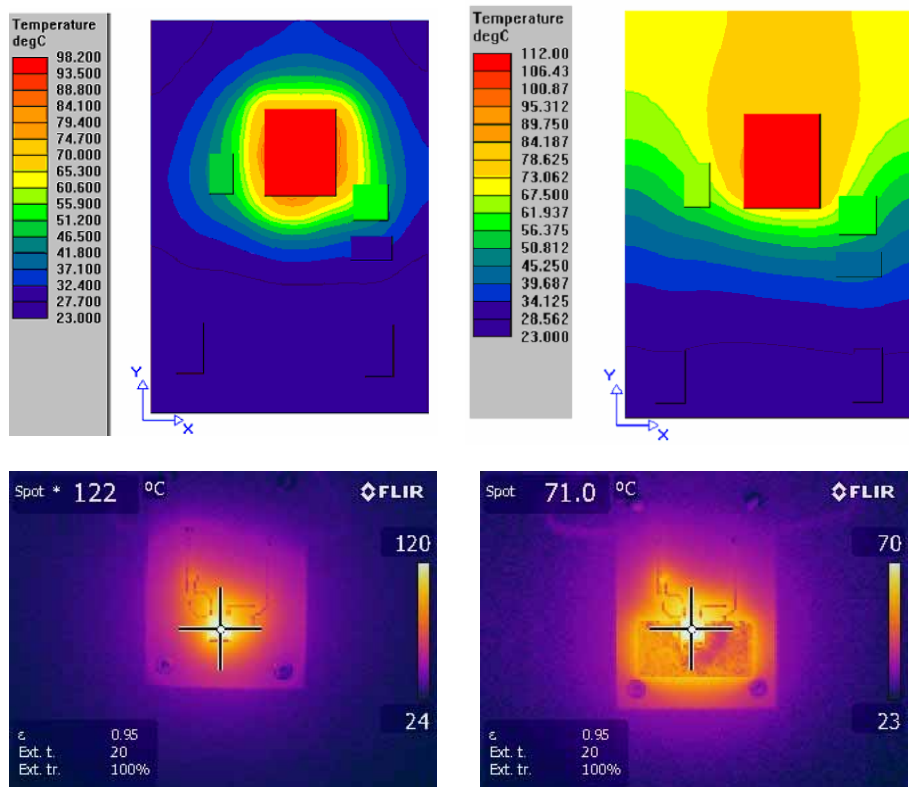
### IV. EXPERIMENTÁLNE POSÚDENIE SIMULÁCIÍ

Pre základné overenie kvality simulácií bolo vytvorených niekoľko typov vzoriek slúžiacich ako modely elektronického zariadenia s jedným alebo niekoľkými zdrojmi tepla. Ako zdroje tepla boli využité výkonové tranzistory v puzdrách TO225AA pre THT a TO263 pre SMT v zapojení prúdovej brzdy. Bolo použitých niekoľko topologických návrhov na substráte FR4 s hrúbkou medenej fólie  $75 \mu\text{m}$  za účelom posúdenia kvality simulačných výstupov v modelových situáciách. Modelovaný bol vplyv chladiacich medených plôch, vplyv použitia pasívnych chladičov na THT súčiastkach a vplyv vzájomného tepelného ovplyvňovania komponentov cez spoločnú medenú plochu (obr.1).



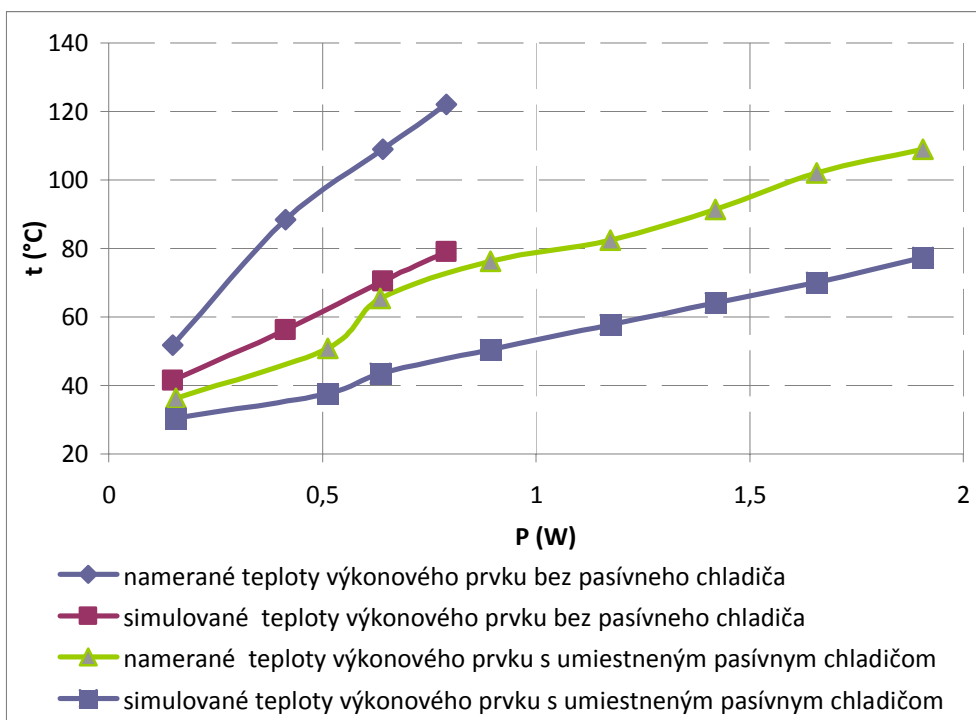
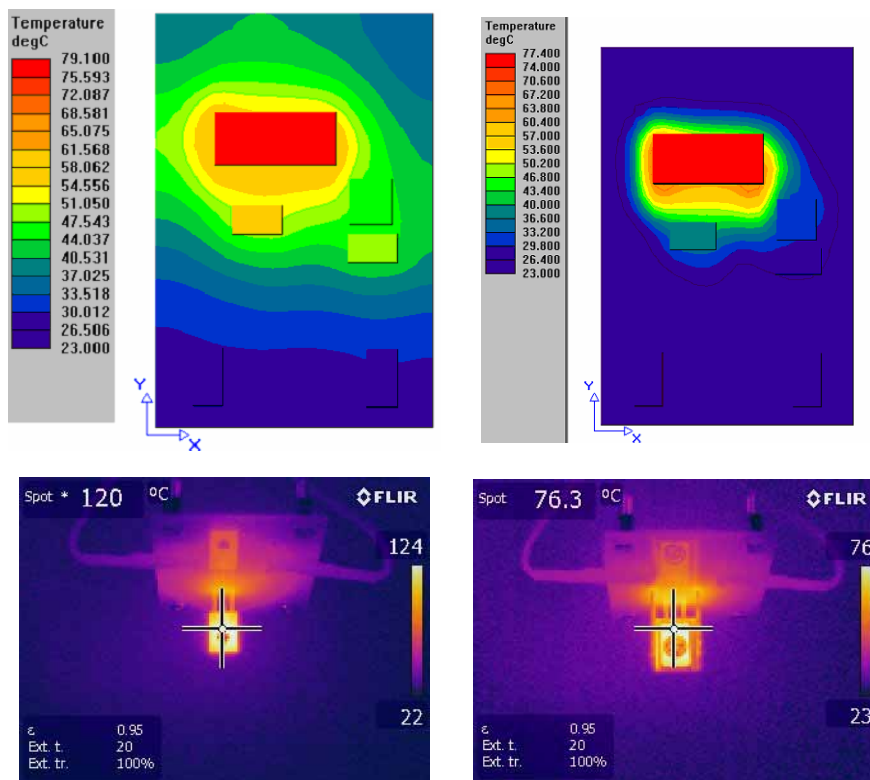
Obr. 1 Realizované experimentálne topológie: a) výkonový SMT komponent bez chladiacej plochy, b) výkonový SMT komponent s medenou chladiacou plochou, c) výkonový THT komponent bez pasívneho chladiča, d) SMT komponenty s modulovateľným výkonom pre analýzu vplyvu vzájomnej tepelnej väzby

Výsledky simulácií softvérom HyperLynx® Thermal boli porovnané s výsledkami merania pomocou termokamery FLIR E30. Hlavným vyhodnocovaným parametrom bola teplota výkonových prvkov, pričom bola posudzovaná aj zhoda tvaru teplotného poľa. Merania boli realizované do maximálnej teploty puzdra výkonových prvkov  $125^{\circ}\text{C}$ . Všetky merania boli uskutočnené bez núteného prietoku pracovnej atmosféry.



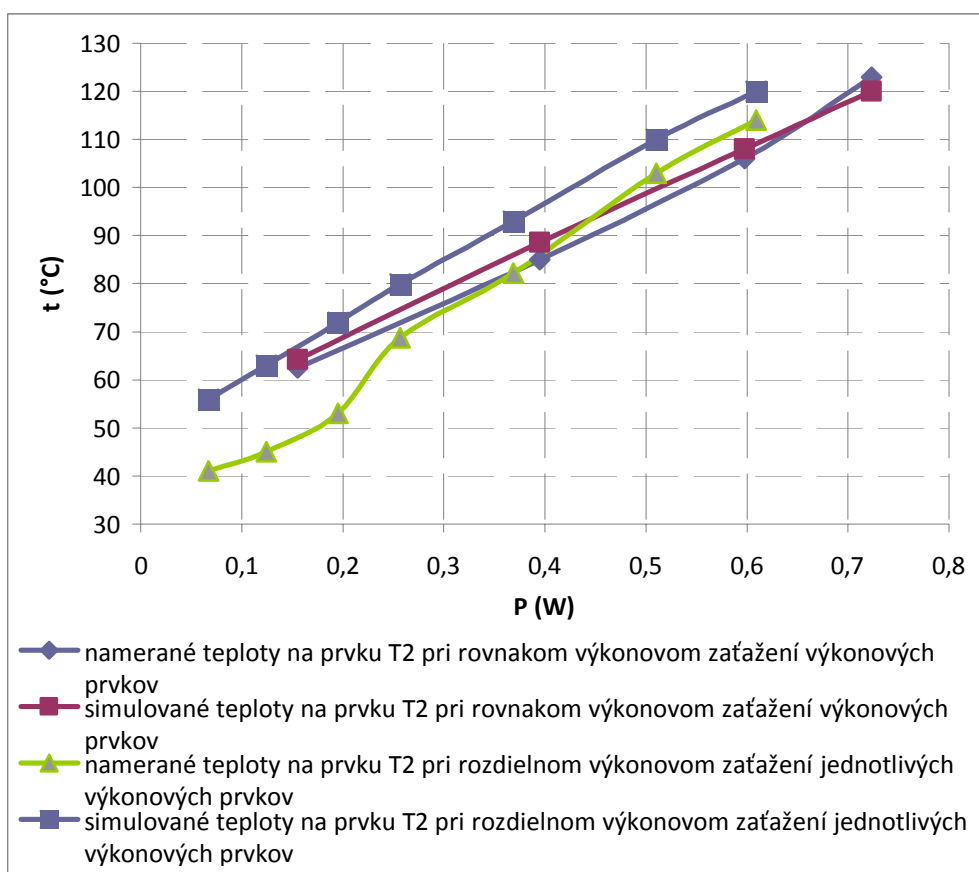
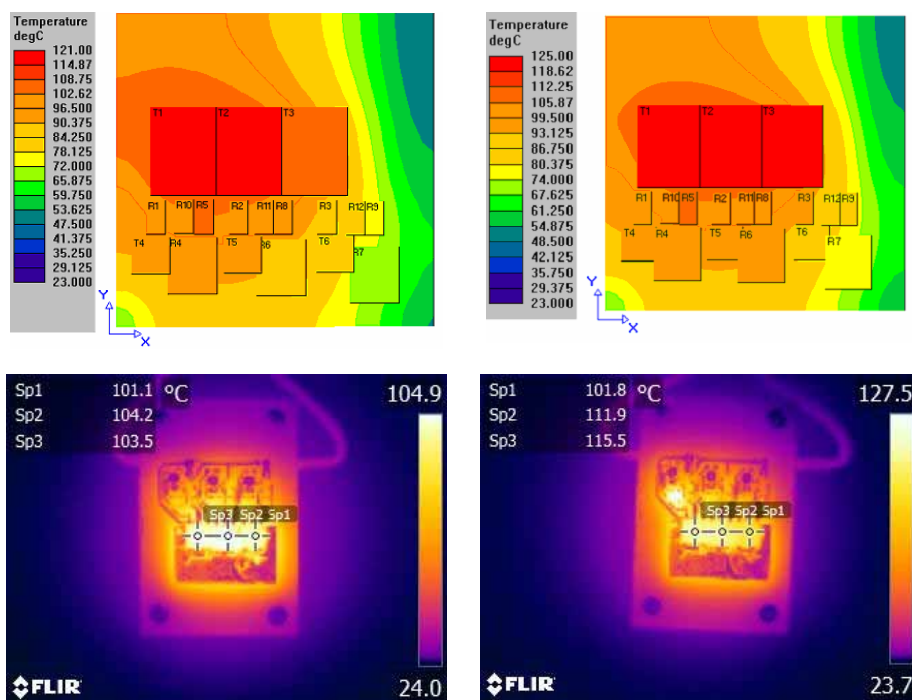
Obr. 2 Výstupy simulácií, termovízne snímky a priebehy nameraných a simulovaných teplôt v závislosti od stratového výkonu pre výkonový SMT prvok bez chladiacej plochy (vľavo) a s medenou chladiacou plochou (vpravo).

Porovnanie výsledkov simulácií s teplotným poľom nameraným termokamerou FLIR E30 poukazuje na fakt, že v prípade výkonového prvku bez chladiacej plochy sa výsledky simulácií rozchádzajú so skutočnými teplotami výkonového prvku až o 20°C, pričom tvar teplotného poľa korešponduje so skutočnosťou. V prípade topológie s chladiacou medenou plochou sú výsledky simulácií vo všetkých smeroch prakticky zhodné so skutočnými pomermi.



Obr. 3 Výstupy simulácií, termovízne snímky a priebehy nameraných a simulovaných teplôt v závislosti od stratového výkonu na výkonovom THT prvku bez pasívneho chladiča (vľavo) a s pasívnym chladičom (vpravo).

Výsledky simulácií sa v tomto prípade výrazne rozchádzajú so skutočným teplotným poľom. V prípade výkonového prvku bez pasívneho chladiča je odchýlka -30 %, v prípade použitia pasívneho chladiča -20 %. Tvar simulovaného teplotného poľa v oboch prípadoch korešponduje so skutočným.



Obr. 4 Výstupy simulácií, termovízne snímky a priebehy nameraných a simulovaných teplôt v závislosti od stratového výkonu pre výkonový SMT prvok bez chladiacej plochy (vľavo) a s medenou chladiacou plochou (vpravo).

V prípade rovnakej výkonovej straty na všetkých výkonových prvkoch výsledky simulácií veľmi dobre korelujú so skutočnosťou. Pri rozdielnej výkonovej strate jednotlivých prvkov je korelácia v oblasti nižších výkonov horšia, smerom k vyšším teplotám sa odchýlka znižuje. Pri aproximácii nameraných hodnôt odchýlka nepresahuje výrobcom deklarovaných 10 %.

Zo súboru nameraných hodnôt je možné usudzovať, že softvér HyperLynx® Thermal dosahuje dobrú zhodu pri použití SMT komponentov a relatívne veľkú odchýlku pri použití THT komponentov (obr. 3 a 4). Odchýlky sú pravdepodobne spôsobené nedostatočne optimalizovanými modelmi súčiastok, resp. jednostrannou optimalizáciou týchto modelov pre



prostredie s núteným prúdením atmosféry, čomu nasvedčuje aj mierna pozorovateľná deformácia teplotných polí aj v prípade nastavenia nulového prúdenia vzduchu.



#### V. ZÁVER

Softvérový nástroj HyperLynx® Thermal predstavuje užitočnú pomôcku pri návrhu moderných miniaturizovaných systémov. Aj keď z realizovaných overovacích simulácií vyplynulo, že pri použití THT komponentov je jeho presnosť nevyhovujúca, pri simulácii DPS s SMD komponentmi, ktoré v súčasnosti predstavujú najpoužívanejší typ z hľadiska druhu montáže sú výsledky simulácií pomerne presné (resp. vo výrobcovi udávanej tolerancii). Rovnako dobré výsledky dáva tento softvér pri simulácii vzájomného teplotného ovplyvňovania SMD komponentov, ktoré sú umiestnené v tesnej blízkosti. Kým chladenie THT komponentov nepredstavuje v praxi obtiažne riešiteľný problém, resp. jedná sa o problém riešiteľný aj dodatočne použitím chladiča s menším tepelným odporom, prehrievanie SMD komponentov je bez podstatných konštrukčných zmien riešiteľné len výnimočne.

Práve v tejto oblasti je s výhodou uplatniteľný softvér HyperLynx® Thermal, ktorý umožňuje teplotnú simuláciu CAD návrhu a dáva možnosť včasnej identifikácie problémových oblastí, vrátane skúmania možností ich eliminácie.

Vďaka skutočnosti, že problémové detaily sú identifikované už v procese CAD návrhu, je možné jednoducho vykonať konštrukčné zmeny a potenciálne problémy eliminovať, čo má priaznivý dopad na celkové vývojové náklady.

#### POĎAKOVANIE

	Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.
	Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ

#### LITERATÚRA

- [1] ELBERG, S. – MATHONNET, P.: Odvod tepla z elektronických zařízení. Praha, SNTL 1983, ISBN 04-533-83
- [2] Datasheet HyperLynx® Thermal. [cit 2014-2-5] Dostupné na internete: [http://s3.mentor.com/public\\_documents/datasheet/pcb/simulate-thermal-ds.pdf](http://s3.mentor.com/public_documents/datasheet/pcb/simulate-thermal-ds.pdf)
- [3] FLIR E30, Uživatelská příručka
- [4] JAN LÍBAL: Využití termovízní kamery pro měření teploty elektronických stroju.
- [5] Brno: VUT, 2009. [cit 2013-10-15]. Dostupné na internete: [https://www.vutbr.cz/www\\_base/zav\\_prace\\_soubor\\_verejne.php?file\\_id=14118](https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=14118).

# Prototyp palubného počítača na báze OBD2 s bezdrôtovým rozhraním

*Pavol CABÚK, Slavomír KARDOŠ*

Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky, Technická  
univerzita v Košiciach, Slovenská republika

pavol.cabuk@tuke.sk, slavomir.kardos@tuke.sk

**Abstrakt** — Predmetom príspevku je konštrukčný návrh a praktická realizácia zariadenia s funkcionalitou palubného počítača pripojiteľného na diagnostický konektor automobilu. Keďže sa jedná o zariadenie, ktoré bude pripojené do konektora diagnostického systému OBD2, bolo potrebné vyriešiť rad problémov týkajúcich sa rozmerov, mechanicky odolného zapuzdrenia a jednoduchej ovládateľnosti spolu s dobrou čitateľnosťou zobrazovaných údajov.

**Kľúčové slová** — Bluetooth, diagnostika automobilov, OBD 2, palubný počítač

## I. ÚVOD

Skutočnosť, že všetky dnešné automobily disponujú diagnostickým rozhraním, umožňuje pomerne jednoduchú konštrukciu zariadenia poskytujúceho funkcionality palubného počítača, druhej palubnej dosky (pre výcvik v autoškole), ako aj základnej diagnostiky.

Jednou z dôležitých funkcií palubného počítača je zobrazenie aktuálnej alebo priemernej spotreby paliva. So spotrebou paliva súvisia mnohé parametre, ktoré dokážu vodičovi napomôcť k jej zníženiu alebo zefektívneniu celkového chodu vozidla. Závaž motora spolu s polohou škrtiacej klapky patria medzi také parametre, ktoré informujú vodiča o aktuálnych jazdných podmienkach, a tak nepriamo vplyvajú na spotrebu paliva. Ďalšie parametre ako teplota chladiacej kvapaliny, teplota nasávaného vzduchu a mnohé iné, ktoré sú závislé od vybavenia automobilu, pomáhajú v rôznych situáciách, kedy vodič o nich nemá žiadne informácie alebo sú k dispozícii s veľkým skreslením a oneskorením.

## II. DIAGNOSTICKÉ ROZHRAŇIA AUTOMOBILOV

V 50-tych rokoch 20. storočia sa diagnostika vykonávala len pomocou jednoduchých meraní a skúsený mechanik musel problém odhaliť svojimi zmyslami. Vylepšovaním technológií ako v automobiloch, tak aj v meracej technike sa v 60-tych rokoch začali do automobilov zabudovávať senzory, ktorých charakteristiky sa dali merať a následne porovnávať so štandardizovanými hodnotami.

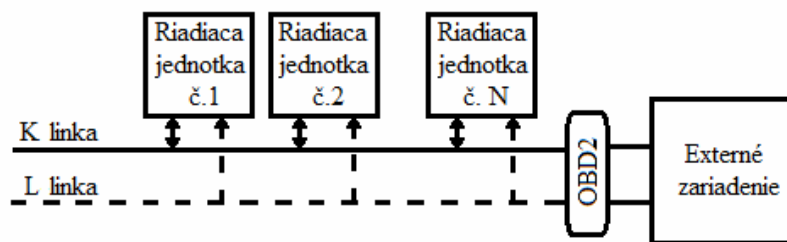
Prvý elektronický systém schopný vyhodnocovať poruchy bol uvedený v Európe v roku 1969 [1]. Organizácia SAE (Society of Automotive Engineers) predstavila v roku 1988 sadu testovacích signálov a odporúčaný typ diagnostického konektora. V roku 1994 vstúpilo do platnosti nariadenie, aby každý automobil vyrobený v Kalifornii od roku 1996 bol vybavený OBD2 štandardom, do ktorého boli implementované odporúčania SAE. V krajinách Európskej únie bol v roku 1998 schválený zákon nariaďujúci aby automobily s benzínovými motormi zakomponovali do výbavy EOBD (European On-Board Diagnostics) od roku 2000 a pre automobily s dieselovými motormi s platnosťou od roku 2004. EOBD je technickými parametrami zhodná s OBD2.

Moderné automobily obsahujú veľké množstvo riadiacich komponentov, ktoré medzi sebou vyžadujú spoľahlivú a rýchlu komunikáciu. CAN (Controller Area Network ISO11898) takúto komunikáciu umožňuje, preto je od roku 2008 každý automobil uvedený do prevádzky v USA povinný podporovať komunikačný štandard ISO 15765-4, ktorý je variantom CAN [2].

## III. KOMUNIKAČNÁ ZBERNICA

Komunikácia prebieha po zbernici, zloženej z K linky (kolík č.7) a voliteľnej L linky (kolík

č.15). K linka je obojsmerná komunikačná linka, po ktorej sú vysielané riadiace signály z externého zariadenia a dáta vyslané komponentmi pripojenými na zbernicu. Ak L linka absentuje, K linka môže byť použitá na inicializáciu komunikácie na zbernici. Hlavnou úlohou jednosmernej komunikačnej L linky je vyslanie inicializačnej sekvencie. L linka môže taktiež prenášať riadiace signály, avšak len smerom z externého zariadenia na zbernicu automobilu. Po inicializácii zostáva L linka vo vysokej logickej úrovni.[5]



Obr. 1 Zbernica komunikačného protokolu ISO 9141-2 a ISO 14230-4 [5]

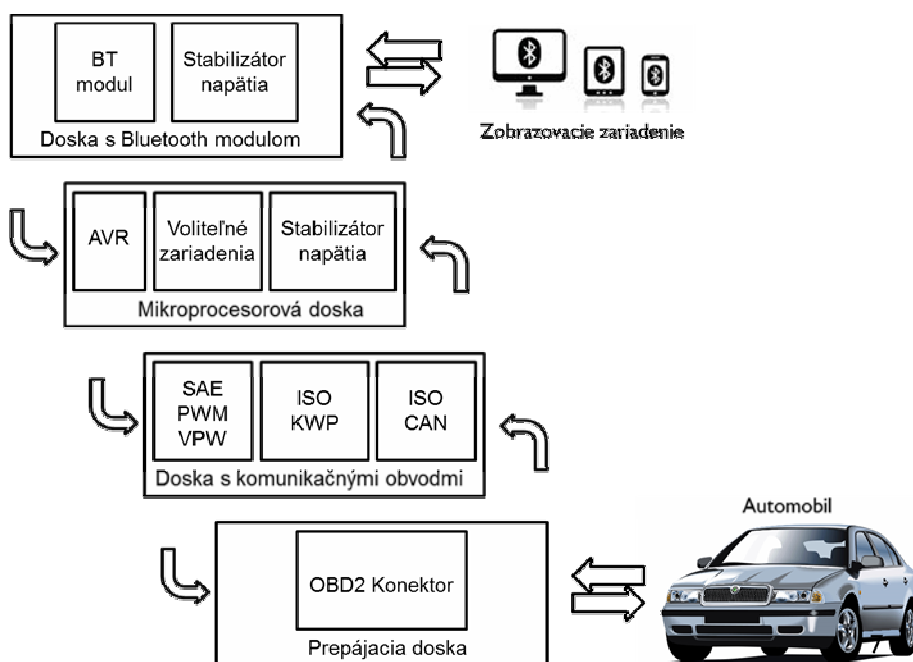
#### IV. KONCEPCIA PALUBNÉHO POČÍTAČA

Palubný počítač v automobile musí obsahovať zobrazovaciu jednotku, ktorá je čitateľná aj pri zhoršených svetelných podmienkach. Ovládacie prvky palubného počítača tvoriace rozhranie medzi užívateľom a automobilom musia byť dostatočne robustné a umožňovať presné, rýchle a intuitívne ovládanie bez zbytočného odvádzania pozornosti vodiča.

Zobrazované údaje musia mať dostatočnú výpovednú hodnotu. Preto palubný počítač vozidla zobrazuje hodnoty, ktoré vodičovi prístroje na palubnej doske neposkytujú alebo sú nevhodne umiestnené a je na ne zhoršený výhľad, či už pre výšku vodiča alebo neergonomicky navrhnutý interiér vozidla.

#### V. KONŠTRUKCIA PALUBNÉHO POČÍTAČA

Návrhom palubného počítača s Bluetooth modulom bolo eliminovaných množstvo problémov vyskytujúcich sa pri návrhu palubného počítača s pevným LCD displejom. Zobrazovaciu jednotku takéhoto zariadenia môže tvoriť akékoľvek zariadenie s podporou bezdrôtovej technológie Bluetooth. Vďaka plneprogramovateľnému mikrokontroléru ATmega324PA s množstvom periférií je možné vytvoriť základ palubného počítača pre akékoľvek vozidlo s možnosťou ľubovoľnej interpretácie získavaných dát.



Obr. 2 Bloková schéma modulárnej konštrukcie palubného počítača.

Minimalizácia rozmerov celkového zariadenia bola docielená návrhom viacerých menších dosiek naskladaných nad sebou. Výhodou tejto koncepcie je možnosť budúcich vylepšení alebo

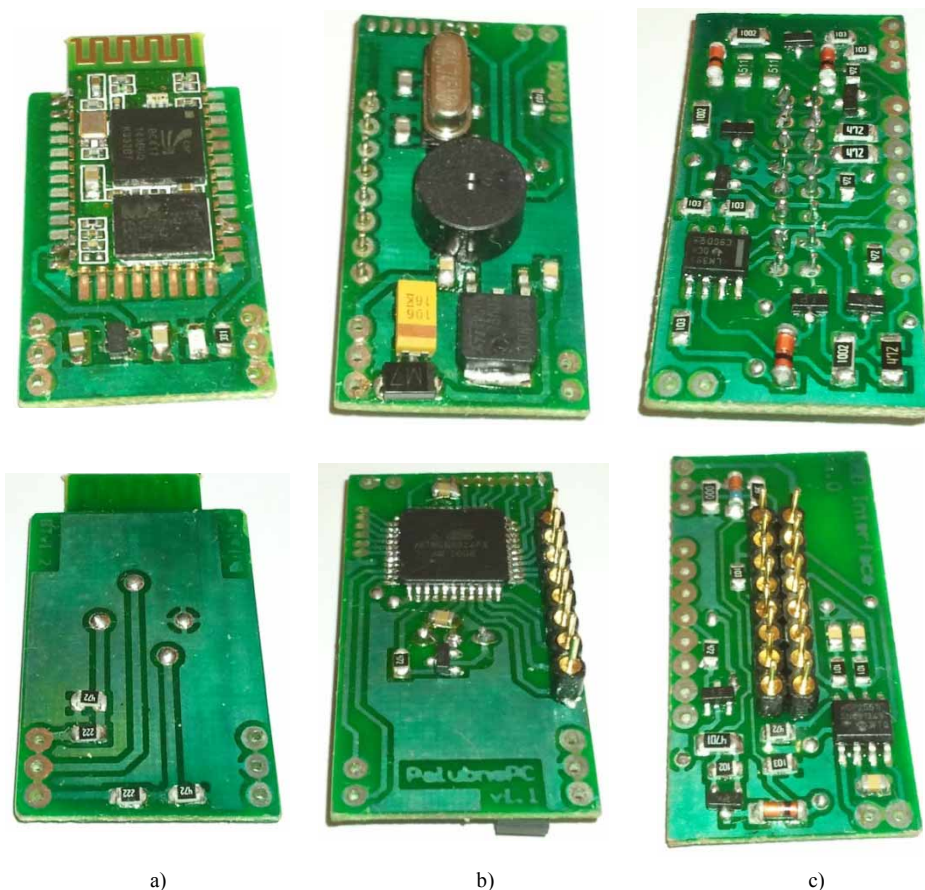
modifikácie jednotlivých obvodových dosiek nezávisle na ostatných.

Prepájacia doska predstavuje vstupný prvok tvoriaci prepojenie medzi diagnostickým konektorom vozidla a jednotlivými doskami palubného počítača. Doska s komunikačnými obvodmi obsahuje rozhrania pre komunikačné protokoly bežne používané v automobiloch s diagnostickým konektorom OBD2.

Základná výpočtová jednotka palubného počítača je tvorená mikroprocesorovou doskou, kde sa spracovávajú, uchovávajú a vyhodnocujú zaznamenané parametre automobilu. Mikroprocesorová doska obsahuje mikrokontrolér ATmega324PA s externým zdrojom časovania tvoreným kryštálom. Mikrokontrolér je programovateľný priamo v obvode pomocou ISP programátora cez SPI rozhranie, prípadne bezdrôtovo.

Komunikačná doska s Bluetooth modulom sa používa na komunikáciu s externými zariadeniami, ktoré dokážu interpretovať údaje vysielané z palubného počítača na svojej zobrazovacej jednotke. Externými zariadeniami môžu byť všetky zariadenia podporujúce rozhranie Bluetooth a sú schopné prijímať a vysielat' potrebné príkazy kompatibilné s OBD2 komunikačnými protokolmi. Pri použití dosky s Bluetooth modulom sa tak navrhnutý palubný počítač správa ako správca komunikačnej zbernice automobilu, kde si užívateľ na externom zariadení zvolí požadované údaje na zobrazenie.

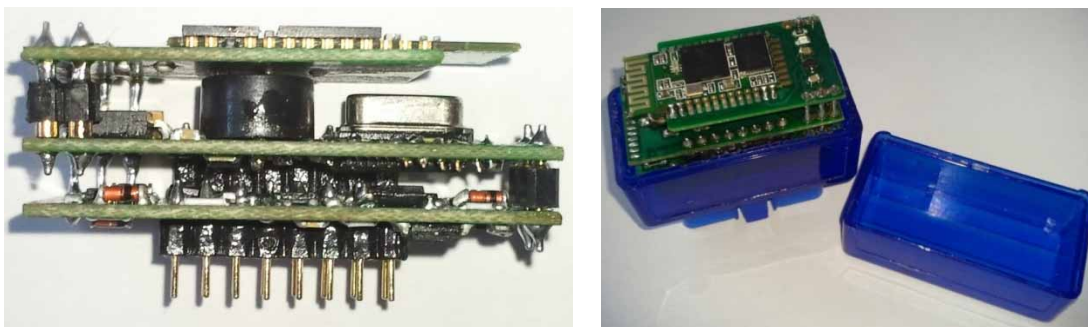
Elektrické a mechanické spojenie jednotlivých dosiek je realizované vhodne rozmiestnenými precíznymi dutinkovými lištami alebo nerozoberateľne prispájkovanou kolíkovou lištou.



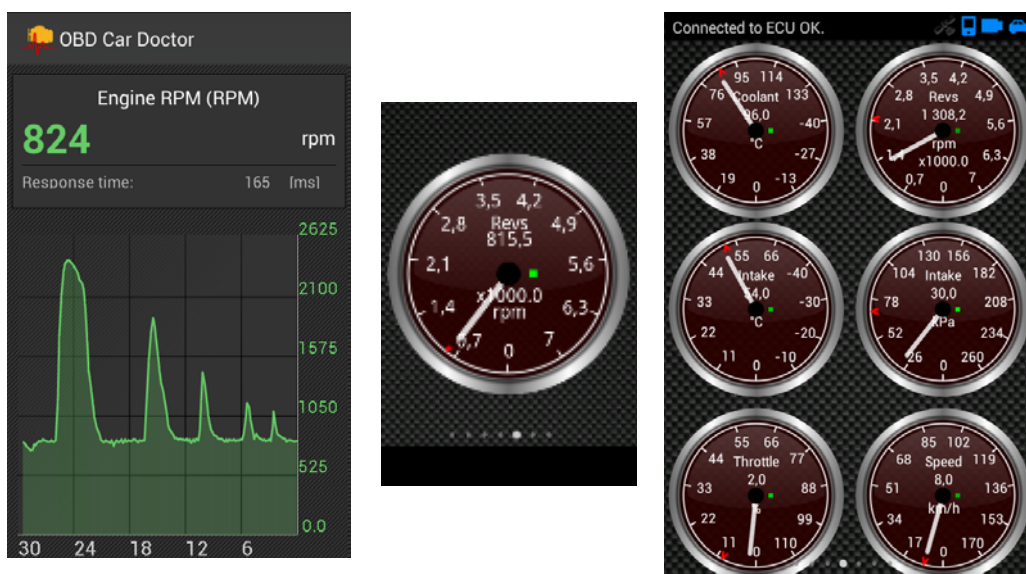
Obr. 3 Osadené obvodové dosky palubného počítača: a) komunikačná doska s bluetooth modulom, b) mikroprocesorová doska, c) doska s komunikačnými obvodmi

Prispájkovaním prepájacej dosky k vývodom puzdra vzniká stabilná základňa pre pripájanie obvodových dosiek palubného počítača.

Pre využitie funkcionalít hardvéru je potrebné zobrazovacie zariadenie s nainštalovanou aplikáciou pre interpretáciu prijímaných parametrov z automobilu (napr. OBD Car Doctor) bezdrôtovo pripojiť s palubným počítačom.



Obr. 4 Prepojené obvodové dosky palubného počítača a palubný počítač umiestnený v puzdre



Obr. 5 Možnosti zobrazenia snímaných údajov z palubného počítača v rôznych aplikáciách

## VI. ZÁVER

Prototyp palubného počítača na báze OBD2 je súčasťou systému na diagnostiku a zobrazovanie parametrov automobilu. Systém tvorí palubný počítač s Bluetooth modulom a kompatibilné mobilné zariadenie podporujúce rovnakú bezdrôtovú technológiu prenosu dát. Zariadenie interpretuje prijímané parametre z automobilu na dotykovom grafickom displeji vo forme grafov, analógových ciferníkov alebo ako číselnú hodnotou. Takýto palubný počítač tvorí finančne výhodnejšiu náhradu originálneho zariadenia v automobile, umožňuje využívanie „druhej palubnej dosky“ pri výcviku vodičov, prípadne môže (v obmedzenej miere) doplniť výbavu staršieho automobilu s kompatibilným komunikačným protokolom diagnostického systému OBD2.

## POĎAKOVANIE

	Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.
	Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ

## LITERATÚRA

- [1] VW's pre-service diagnosis. In: Commercial Motor. 17.10.1969, s. 42. Online: <<http://archive.commercialmotor.com/article/17th-october-1969/42/vws-pre-service-diagnosis>>
- [2] Environmental Protection Agency :Control of Air Pollution From New Motor Vehicles and New Motor Vehicle Engines, Publikované: 20.12.2005 s. 75403-75411. Online: <<http://www.epa.gov/fedrgstr/EPA-AIR/2005/December/Day-20/a23669.htm>>
- [3] SAE J1962:2002, Diagnostic Connector
- [4] ISO 15765-4:2005, Road vehicles - Diagnostics on Controller Area Networks (CAN)
- [5] Volkswagen, Group of America : K-line Communciation Description. [online] 2010, Online: <[http://www.obdclearinghouse.com/index.php?body=get\\_file&id=1343](http://www.obdclearinghouse.com/index.php?body=get_file&id=1343)>



# Design of a Wireless Sensor Network for the Surface Water Quality Monitoring

Ludmila MACEKOVÁ

Department of Electronics and Multimedia Communications, Faculty of Electrical  
Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>ludmila.macekova@tuke.sk

**Abstract** — In this paper, main aspects of the design of the water quality monitoring wireless sensor network (WSN) are described. Selecting of a few main water quality indicators, principles of sensors and a design of the architecture of such network are introduced. Relations of such idea and activity to both the Slovak and the European legislatives are mentioned. This WSN is intended to monitor continuously the ecological status of the river, to participate in the early warning system and to contribute to the environmental protection.

**Keywords** — water quality indicator, sensor, water quality probe, wireless sensor network

## I. INTRODUCTION

Generally, the water quality has very important influence on the life of human, animals and herbage. The presence of different matters dissolved or scattered in the water affects the physical, chemical and biological properties of the water (for example acid-base equilibrium - pH parameter, conductivity, hardness, colour, opacity or cloud, smell, and temperature, and sometimes some foam on the water surface, and so on). Acid rainfall, wastewater, land use and a different other human activity can change the water properties and destroy the life conditions in the surface water bodies.

More specific scientific branches deal with these biological, chemical, ecological, legislative, etc. problems and with protection of life in all types of the water bodies and their surroundings. The institutions as the SHMI (Slovak Hydro-Meteorological Institute), the WRI (Water Research Institute – Výskumný ústav vodného hospodárstva), the SWME (Slovak Water Management Enterprise - Slovenský vodohospodársky podnik) and its relevant branch river basins managements (e.g. Povodie Bodrogu a Hornádu, etc.) collaborate on the government and legislative level in the Slovakia. They implement also the directives of the European Parliament and of the Council of The European Union [3], [4] in their documents as [5], [6], [7]. Very heterogeneous types of pollution, a lot of the water quality indicators, their dangerous limits, the reference certified measurement methods are listed in these documents, also, the tasks for all participants and responsible members can be found there.

The aim of the regular measurements and evaluation of indicators is to monitor the water-body status and subsequently launch the needed steps with more purposes (an alarm, and so on).

In the frame of the Hungary - Slovak cross-border cooperation programme (HUSK; see below) the idea of the monitoring of the Danube River water quality parameters arose. The aim is to design the wireless sensor network (WSN) for such monitoring and to contribute to the environment protection. The gained outcomes can serve to next applications in the other water streams and water bodies.

Based on the legislative literature, consultations with the specialists, producers literature ([13], [14], etc.), and other references we have decided to monitor only a few main indicators of the water quality. The values of these indicators and their change closely correlate with the other quality parameters from a very long list.

The measurement of the selected parameters by special probes, the calibration and the maintenance of the probes are complicated processes enough. Therefore the use of professional probes, their accessories and communications systems were chosen for the purpose of the WSN

for the water quality monitoring in the water streams.

In the next parts of the paper we will familiarize with the main water quality parameters and their measurement principles in sensors. Then the professional complex equipment communication protocols, the methods of measurements, and, in the end, our conceptual design of the surface water quality WSN will be described.

## II. PRINCIPLES OF MAIN INDICATORS OF THE SURFACE WATER QUALITY AND THEIR MEASUREMENTS

Measurement principles and also limits of the selected main indicators established for the good ecological quality water for the Danube River by [5] are here described. The ecological “good quality” status of the river water is the 2<sup>nd</sup> degree of the 5 degrees scale in the ecological evaluation of the water streams quality [23].

Point of view for the followed explanations is mostly rather chemical than electrical one. Electrochemical processes in glass probes systems with semipermeable membranes (or without them) convert the value of water quality indicator into an electronic signal. The measurement process, however, is much more complex: it contains also the calibration, the using of calibrating solutions (buffers), the measurement of the ambient temperature, and very careful cleaning, maintenance and storage of the probes.

Today’s ambition of developers and producers is to use also the biochemical property vs. optoelectronic information/data relation. These optical methods use luminescent sensing dyes. Optical sensors, working with spectrometry principle are more comfortable and demand less calibration and maintenance. They are often used in the parallelism or before the more precise electrochemical methods in a laboratory but also for a food control, etc.

One more principle is investigated and (e.g. in [20]) also applied: the operation of the Ion-sensitive Field Effect Transistors (ISFETs), like the special class of the Chemical-sensitive FETs (ChemFETs). In the ISFET, the analysed solution is in the direct contact with the gate insulator layer (SiO<sub>2</sub>) over an electroactive membrane, and a metal reference electrode is immersed in the measured solution and operates as the metal gate (i.e. the electrode to control the conducting channel) [20]. For example the Microsens produces ISFET sensors [21] and the Hach Series Meters and Probes apply this phenomena in their non-glass pH probes and multimeters [22].

### A. pH-parameter

In the chemistry, briefly, the pH is a measure of the acidity or basicity of an aqueous solution, which correlates with the hydrogen ion H<sup>+</sup> concentration, expressed in the amount of the moles per litre, noted symbolically by the [H<sup>+</sup>]. The theoretical definition of the pH parameter says it is the negative common logarithm of the hydrogen ion concentration [1], [2]:

$$\text{pH} = -\log[\text{H}^+] \quad (\text{without units}). \quad (1)$$

Some illustrative values of the pH are introduced in the Table 1.

Table 1  
 The pH values of some water solutions [1]

Type of solution	pH (-)	Note
pure fresh water	about 7	pH neutral; at 25°C
drinking water	<6.5, 8.5>	
sea water	close to 8	
stream water	<6, 9>	
acidic solution	< 7	
sour lemon juice	about 2	
strong cleaner solution	>13	
basic solution	>7	alcaline solution

We must accent, that the value of pH closely depends on the temperature of the solution.

The more practical definition of the pH is related to the electrometric mechanism of the pH measurements. Here, the pH is proportional to the potential difference (or the measurable voltage) E in millivolts between both the reference and the indicating electrodes immersed in the measured water solution [16]):

$$E = E_{ind} - E_{ref} = E'_T + K_g(T/F) \ln[H^+], \quad (2)$$

where  $E_{ind}$  is the voltage of the indicator electrode (mV),  $E_{ref}$  is the voltage of the reference electrode (mV),  $E'_T$  is the temperature dependent constant in mV,  $K_g$  is the gas constant (8.3144 J/K),  $T$  is the absolute temperature (K), and  $F$  is the Faraday's constant (96485 C mol<sup>-1</sup>).

Buffers used for the calibration of pH probes are the stable solutions with exactly defined pHs at the defined temperatures.

### B. Conductivity

The conductivity ( $CD$ )– more exactly, the specific conductance – in the S/m units - is defined as:

$$CD = \frac{K_{cell}}{R} \frac{1}{1 + \left(\frac{\alpha}{100}\right) * (T - 25)}, \quad (3)$$

where the  $K_{cell}$  is the cell constant in m<sup>-1</sup> (it equals to  $l/A$ , where the  $l$  is the distance between the electrodes in the galvanic cell and the  $A$  is the area of the electrodes),  $R$  is the measured resistance in  $\Omega$ ,  $\alpha$  is the temperature compensation factor in % per °C, and  $T$  is the temperature in °C [15].

Generally, the measuring of the conductivity means the measuring the ionic (or the ionisable solutes) content of a solution [17]. Mostly the  $CD$  is expressed in the smaller units (mS/m or  $\mu$ S/cm, etc., see Table 2).

Table 2  
 Some selected values of the water conductivity [17]

Type of the water	CD	Temperature	Note
Danube River	< 70 mS/m	< 23 °C	for the 2 <sup>nd</sup> degree of the quality [5]
drinking water	0.5 – 1 mS/cm	25 °C	
pure water	0.05 $\mu$ S/cm	25 °C	

The conductivity probe provides a current or a voltage output, or the frequency value output of the converter. Input impedance of the measurement circuit must be high. The accuracy and the resolution of the conductivity measurements depend on the contamination of the sensor electrode surface, electrochemical effects, and on the temperature dependency of the sensor.

Two basic types of sensors for conductivity measurements with different properties are the *contacting* (conductive) sensors and the electrode-less (*inductive*) sensors.

### C. Dissolved Oxygen

Presence of the oxygen is one of the necessary conditions of life. The level of the dissolved oxygen (DO) is important in the lakes, streams, fish farming, and also in waste water treatment. For example, more than 7 mg of the oxygen per litre must be dissolved in the water for the good quality status of the Danube River (the 2<sup>nd</sup> degree) [5]. Less than 5 mg/l of the oxygen in the water causes the deterioration of the health of the fish, less than 2 mg/l causes their death.

Content of DO in the water closely depends on the water temperature (it decreases with increasing of the temperature).

The principle of the DO measurements can be again an electro chemical process in *polargraphic* or *galvanic cells*, or an *luminescent method* in LDO (Luminescent Dissolved Oxygen) probes [18]. The dissolved oxygen sensors with the optical principles measure the interaction between the oxygen and certain luminescent dyes [19].

### D. Redox parameter or ORP

The REDOX (reduction – oxidation) parameter or the ORP (the oxidation-reduction potential) of the solution expresses its tendency to oxidize or reduce another chemical substance [11]. The oxidation is the loss of electrons, and the reduction is the gain of electrons by an atom, molecule, or ion.

ORP can be calculated by the Nernst equation, and evaluated and measured in millivolts between about -1500 and + 1500 mV.

The REDOX ability of a solution depends on the temperature [11] and very closely relates to the life in the water. A decreasing redox potential indicates the deterioration of the water quality. The ORP changes after appearing of a new species or after changing of the concentration of an existing species. The ORP monitoring and controlling can early indicate an



attendance of the cyanide, chlorine or chrome in the surface water, some destroying microorganisms in the swimming pool using the chlorination, etc. For example the +680 mV value of the ORP is recommended for the pool water.

The pH control is mostly realized simultaneously with the ORP measurement.

An ORP sensor consists of an ORP electrode and a reference electrode, and uses the potentiometric principle like the pH measurements.

#### *E. Temperature*

The temperature affects many physical, biological, and chemical characteristics of the water. As it was mentioned above the temperature influences the amount of the oxygen dissolved in the water, the rate of the photosynthesis of water plants, the metabolic rate of animals, etc. The thermal pollution means the increasing of the water temperature, which can be of the various origin – waters from warmed urban surfaces, waste waters, a removing of trees from river banks, etc. By [5] for the good state (the 2<sup>nd</sup> degree) of Danube River, the temperature of the water must not be greater than 23°C in the upper section of the stream on the Slovak territory. Each type of the surface water has its own temperature limit.

Resistance principle of thermometer is mostly used in the water quality measurements. This means there are the sensors with some distinguish temperature depending resistor. Resistance rather than voltage is measured.

In the resistance thermometers the resistance of a metal (mostly the platinum) caused by the “defects” disturbing movement of the electrons is utilized. This property is stronger with arising temperature. The products as PT100, PT500, PT1000 based on the platinum are the reliable professional sensors, which were tested and utilised in many applications [9]. They allow to measure the temperature in the range of -200 up to +850 °C with the linear characteristic.

Thermometers can be used for the purpose of temperature measurements, but also for the temperature compensation of other indicator measurements. The thermometers are, therefore, unavoidable component parts of water quality multimeters and probe systems.

### III. PROFESSIONAL MEASUREMENT SYSTEMS

Some high specialized producers offer a lot of the less or more robust and suitable probes or probe systems, mentioned briefly in the next text.

#### *A. Multimeters*

Several hand-held measurement devices (e.g. [24]) which are connectable with more types of probes are known. They can measure and/or recalculate several parameters and have further functionality too. They can serve for combined measurement methods (without and with a man attendant), and also for the calibration and the reference measurements. Multi-parameter multi-channel devices can measure a few parameters simultaneously. They can generate the alarm at the exceeding the defined limits, or in the case, when it is needed to calibrate or maintenance some probe.

#### *B. Multiprobe systems*

Today more producers in the area of water quality measurements offer not only freestanding probes but also so called multiprobe systems, or multiparameter probes. The multiprobe system (e.g. [14]) has the common container, power supply and communication interface for all optional probes (3 or more than 10 ones) for several water quality indicators measurements. The multiprobe can have its own battery, the cleaning brush with the automatic function, the protective cage and the outgoing connecting cable for the power supply and communications. These over-elaborated systems can be fastened on the buoy, and can operate continuously a few weeks without maintenance.

### IV. COMMUNICATIONS PROTOCOLS

The RS-232 or SDI-12 communications protocols can be chosen for the communications (setting, operational control and data logging) between a data-logger and water quality sensor.

RS-232 standard is the known serial system and interface designed to the communications between a PC and other data communicating devices and peripherals. It was suitable also to the telephone lines transmissions. This system is considered the good processed and completely collision-less physical layer of communications. It is however displaced by USB interfaces in modern PCs.

SDI-12 [8] also stands for the serial data interface and communications protocol. This

standard was established for the interconnecting and communications between data recorders and micro-processor based sensors (intelligent sensors). SDI-12 provides the serial data speed up to 1200 baud. This system requires the minimal current drain, can use one data recorder with multiple sensors on a single cable and allows the sensors to make the complex self-calibration.

## V. METHODS OF MONITORING OF WATER QUALITY IN THE WATER STREAMS

### A. State-of-the-art

Water quality monitoring using the remote sensing by the optical sensors and a satellite system in support of the Europe Union Water Framework Directive was described in [26].

The network of monitoring stations system is described in [25]. The waste waters in the Austria and Danube River was monitored by modular water quality monitoring stations, equipped with powerful industrial PCs for the sensors control and data storage. Sensors were continuously immersed into the liquid and directly linked to the PC via a bus-interface or via an analogue input. Data from the stations were automatically processed and stored in the central database.

Authors of the [9] report about automatic stations for the monitoring the river water quality in Serbia. They measured more parameters in the Tisa, Sava, Kolubara and Danube rivers and planned to design the early warning system based on these stations. They used multiprobe systems (by YSI [13]), saved the data using data-loggers, transmitted them through the GSM modem, and stored them in the central data-base. Results were/are accessed by Serbia RHMS server (Republik Hydrometeorological Service). The sensors for the pH, DO, CD, surface-level and temperature measurements were based on the more different measurement principles. An industrial PC (touch panel) with the software for the measurement, storing, display and data sending was installed in each automatic station. Wireless and optical fiber links were used to data transferring and control.

Xylem Analytics Manufacturer in [27] offers buoys like floating platforms with water quality monitoring probes systems (by YSI [13]) and the remote telemetry system.

Now the functionless automatic station is located at Slovakia Uh River [12]. In '90s the station automatically sampled water from the water stream, transported the samples into the measurement station on the river bank and sent data to the control center. The next steps were the data evaluation and store, and distribution of the alarm messages via cellular into the national and international early warning system. The surface level, temperature, pH, CD, DO, petroleum products and some others were measured.

### B. Our concept of the design of the wireless sensor network for the water quality monitoring

Idea of our wireless sensor network (WSN) for the water quality monitoring is presented in the Fig.1. The sensor nodes realized by multiprobes can be placed on the buoys, floating on the water surface. Both the buoy and multiprobes measurement platform must be professional manufactured. A data processing unit and a wireless transceiver unit will be on the buoy, except probes systems. This equipment will realize the translation of the measured data into the messages of the chosen communication protocol and send them by the wireless communications link to the center on the river bank, or/and to the remote center with the laboratory. The communications connection with the remote center can be realized by a selected mobile communications network.

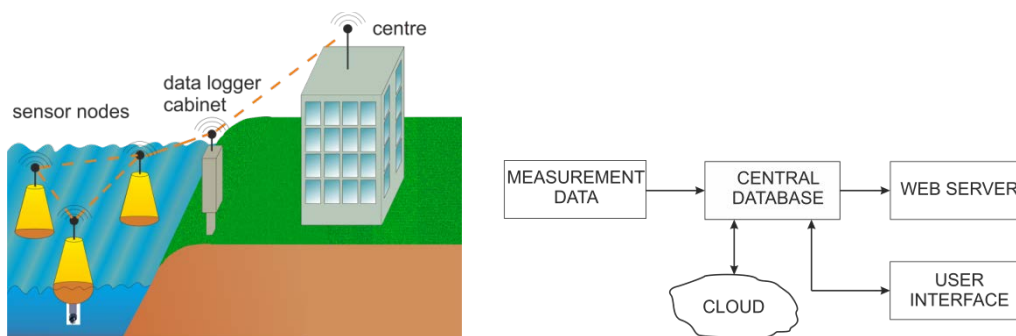


Fig. 1 Illustrations of the measurement architecture of the water quality monitoring WSN

The remote center will be equipped by (except the transceiver unit) the data storing, data processing and analyze hardware. Naturally, communications must include also the link to the water quality laboratory for the manual verifying, analyzing the data and realizing the special

steps at some defined results. Web pages with some selected outgoing data and information will be publicly accessed besides the authorize database and software.

Because of the huge amount of the data, gained by the “continuous” measurements, and because of the requirement to save the long history of measurements the usage of the cloud storage method for storing the data is not excluded.

## VI. CONCLUSION

In this article main facts about water quality measurements were described in relation with the aims of our HUSK project - the design of the WSN for surface water quality monitoring. More of mentioned facts are optional for us (selecting of measured parameters, methods of measurements, type of probes, protocol, etc.). The probably preferred way we start is suggested in the last section.

## ACKNOWLEDGMENT

This work is supported by the project Wireless Sensor Network for wAter QaUality Monitoring (WSN-AQUA) (registration number: HUSK/1101/1.2.1/0091).



We support research activities in Slovakia / Project is co-financed from EU funds. This paper was developed within the Project "Centrum excelentnosti integrovaného výskumu a využitia progresívnych materiálov a technológií v oblasti automobilovej elektroniky", ITMS 26220120055.

Our particular thanks belong to the SHMI, WRI and Slovak Water Management Enterprise – Bodrog and Hornad branch river basins managements for the consultation.

## REFERENCES

- [1] Health Canada: [www.hc-sc.gc.ca/ewh-semt/pubs/water-eau/ph/index-eng.php#n2](http://www.hc-sc.gc.ca/ewh-semt/pubs/water-eau/ph/index-eng.php#n2). Reprinted in 1995.
- [2] Emerson, Process management: Theory and Practice of pH Measurement. PN 44-6033/rev.D. Rosemount Analytical Inc., USA, 2010.
- [3] 2008/105/EC - DIRECTIVE 2008/105/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 December 2008 on environmental quality standards in the field of water policy.
- [4] DIRECTIVE 2000/60/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2000 establishing a framework for Community action in the field of water policy.
- [5] Directive No. 269 of Slovak republic government with requirements for good state of water bodies. 2010.
- [6] Zákon 364 z 13. mája 2004 o vodách a o zmene zákona Slovenskej národnej rady č. 372/1990 Zb. ... .pdf (so called “Water Law of Slovak Republik”).
- [7] Water Plan of the Slovak Republik; Ministry of Environment of the Slovak Republic, 2010.
- [8] SDI-12 Support Group, USA: <http://www.sdi-12.org/>
- [9] S. Mijovic, B. Palmar: Water Quality Monitoring Automation of Rivers in Serbia. Facta Universitatis, Series: Working and Living Environmental Protection Vol. 9, No 1, 2012, pp. 1 10
- [10] Sawi Mess- und Regeltechnik AG: Pt100/Pt1000 platinum resistance thermometer. September 2012.
- [11] Emerson Process Management, Liquid Division, Irvine, CA 92606 USA, © Rosemount Analytical Inc. 2008: FUNDAMENTALS OF ORP MEASUREMENT, Application Data Sheet, ADS 43-014/rev.B, May 2008.
- [12] J. Adamková, D. Ďurkovičová, M. Mikuličková: State of the Uh River is observed by the warning monitoring station in Pinkovce, *Enviromagazin*, 5/2000, p. 10.
- [13] YSI Integrated Systems and Services: <http://www.ysisystems.com/floating-platform-selector.php>.
- [14] HydroTerra Pty Ltd., Australia: <http://www.hydroterra.com.au/products/surface-water-quality-measurement-and-logging>.
- [15] IC Controls, Canada: Technical Notes, The conductivity – Theory and measurement. 2005.
- [16] Radiometer Analytical SAS – a Hach Company Brand: pH theory and practice.
- [17] Van London – Phoenix Co.: Education Series, Conductivity Guide. Glenmont Dr. Houston, USA.
- [18] Hach-Lange, <http://ldo.hach-lange.sk/ldo-sc.php?gclid=CJb-zpmpn78CFaPItAodLycAfA>.
- [19] PyroScience Sensor Technology: [http://www.pyro-science.com/optical\\_oxygen\\_meter\\_firesing\\_O2.html](http://www.pyro-science.com/optical_oxygen_meter_firesing_O2.html). 2014.
- [20] IUPAC publications: Compendium of Analytical Nomenclature. Def. rules 1997.
- [21] Microsens: ISFET, Product Specific Data Sheet [http://www.microsens.ch/products/pdf/MSFET\\_datasheet%20.pdf](http://www.microsens.ch/products/pdf/MSFET_datasheet%20.pdf).
- [22] Hach Company, USA: <http://www.hach.com/ISFET>. 2014.
- [23] I. Bartík, P. Baláži, L. Hamerlík, D. Hlúbiková, K. Kučárová, R. Magulová, J. Makovinská, K. Melová, E. Mišková-Elexová, V. Mužík, Z. Pastuchová, F. Šporka, L. Tóthová, M. Valúchová: Hodnotenie ekol.stavu vodných tokov v SR, časť 1.: Princípy hodnotenia. *Vodohosp. spravodajca* č.5-6/2008, pp.8-10.
- [24] PCE instruments UK Ltd.: Manual PCE PHD 1
- [25] Winkler, S., Pressl, A., Gruber, G., Ecker, M., Fleischmann, N., Kreuzinger, N.: Innovative technology for integrated water quality measurement. Proceedings International IWA Conference, Automation in Water Quality Monitoring AutMoNet 2002, May 21-22, 2002, Vienna, pp. 361-368.
- [26] Q. Chen., Y. Zhang., M. Hallikainen.: Water quality monitoring using remote sensing in support of the EU water framework directive (WFD): A case study in the Gulf of Finland. *Environ Monitoring and Assessment*, 2007, 124, pp.: 157-166.
- [27] Xylem Analytics: <http://www.xylemanalytics.co.uk/productsdetail.php?Harbour-Buoy-Platform-15>.

# Vplyv miesta pripojenia alternatívneho zdroja energie na veľkosť skratového prúdu v distribučnej sústave

<sup>1</sup>Miroslav Kmec, <sup>2</sup>Lubomír Beňa, <sup>3</sup>Lukáš Lisoň

<sup>1,2,3</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>miroslav.kmec@student.tuke.sk, <sup>2</sup>lubomir.bena@tuke.sk, <sup>3</sup>lukas.lison@tuke.sk

**Abstrakt** — V súčasnej dobe je snaha zvyšovať podiel na výrobe elektrickej energie pomocou obnoviteľných (alternatívnych) zdrojov, ktorých činnosť je založená na základe slnečného žiarenia, vetra, vody alebo biomasy. Tieto zdroje elektrickej energie sú šetrnejšie k životnému prostrediu vzhľadom k tradičným tepelným elektrárnam. Tieto zdroje sú v súčasnej dobe vo veľkom počte pripájané do existujúcich distribučných sietí. Vzhľadom k prudkému nárastu týchto zdrojov v distribučných sieťach je potrebné venovať pozornosť vplyvu týchto zdrojov na zmenu skratových pomerov v sieťach. Cieľom tohto príspevku je poukázať, aký vplyv má výber rôznych miest pripojenia novo vybudovaného zdroja energie v našom prípade biomasy na veľkosť skratového prúdu počas poruchy v tejto sieti. Tento článok ďalej popisuje možné variácie pripojenia týchto zdrojov do siete, základné teoretické vedomosti pre výpočet skratov v distribučnej sieti a samotné simulácie namodelované v programe Matlab / Simulink.

**Kľúčové slová** — biomasa, skraty v distribučnej sústave, Matlab, obnoviteľné zdroje energie

## I. ÚVOD

Z pohľadu obnoviteľných zdrojov je elektrárň na biomasu jedna z dôležitých zdrojov pre čiastočné splnenie globálneho dopytu po energii. V rozvojových krajinách sveta, kde povolanie ľudí je prevažne poľnohospodárskeho charakteru s populáciou žijúcou väčšinou v obciach je jeden z najlepších spôsobov ako vyrábať elektrickú energiu a teplo, využitie elektrárne na biomasu. Mnohé z týchto obcí sa nachádza vo veľkej vzdialenosti od väčších miest a ich prepojenie so sieťou je veľmi ťažké, čo má za následok, že ostávajú bez elektrifikácie, alebo majú problémy s nepretržitým napájaním. Pre rozvoj regiónu je nutné využívať energeticky účinných techník a potenciálu obnoviteľných zdrojov energie. Ekonomické riešenie môže byť dosiahnuté správnym hospodárením s energiou, tým že sa využijú existujúce ale aj osvedčené techniky premeny energie, kde tieto prostriedky môžu byť použité pre pokrytie základných potrieb (elektrina, varenia, ohrev vody atď.), čím sa dosiahne zvýšenie sebestačnosti vzhľadom na energetické nároky (potreby). Cieľom je vyrábať elektrickú energiu, produkovať plyn na varenie a iné formy energie na miestnej úrovni s jej efektívnou distribúciou v obci.

Propagáciou novej energetickej politiky vo svete nastal vo veľkom merítku rozmach výstavby obnoviteľných zdrojov energie. Obnoviteľné zdroje energie na báze bioplynu zaznamenali rýchly rast v posledných rokoch aj na Slovensku, kde pripojenie týchto zdrojov do terajších rozsiahlych distribučných sietí môže viesť k značným zmenám ich štruktúry s následnou zmenou skratových pomerov týchto sietí. Preto je naozaj nutné skúmať dopad pripojenia týchto zdrojov do distribučnej siete, vzhľadom na zlepšenie spoľahlivosti dodávky elektrickej energie.

## II. KONŠTRUKCIA ELEKTRÁRNE A VÝROBA ELEKTRICKEJ ENERGIE Z BIOMASY

Elektrárň na biomasu (EB) využíva proces biologickej fermentácie obnoviteľných zdrojov energie na výrobu elektrickej energie a tepla. Ide o kvasenie prírodných odpadov a energetických plodín na výrobu takzvanej zelenej energie. EB majú dôležité miesto v udržateľnej energetickej politike. To je preto, že poskytujú niektoré značné výhody, ktoré iné alternatívne formy energie, ako je solárna a veterná energia nemajú:

- najvyššia energetická účinnosť: bioplyn je prevedený na kombinovanú výrobu tepla a elektrickej energie s účinnosťou 85%,
- nepretržitá prevádzka: anaeróbna fermentácia je biologický proces, ku ktorému dochádza neustále,
- odpad: rôzne typy odpadov môžu byť spracované na energiu,
- flexibilita: odpady z rôznych zdrojov môžu byť spracované spoločne, východiskový bod každého projektu je vždy založený na vstupných surovinách, ktoré sú dostupné v danom regióne. Existuje mnoho surovín, ktoré môžu byť fermentované:
  - plodiny ako je kukurica, repa atď.,
  - hnoj z ošípaných, kurčiat atď.,
  - organické kaly.

Po fermentácii bioplynová stanica vyrába elektrinu, teplo a digestát [1].

### III. PRIPOJENIE ELEKTRÁRNE NA BIOMASU K ELEKTRIZAČNEJ SÚSTAVE

Úroveň napätia, bod pripojenia a metódy pripojenia EB k elektrizačnej sústave určuje príslušný prevádzkovateľ systému na základe technických a ekonomických kritérií.[2]

Napät'ové úrovne pre pripojenie:

V závislosti od inštalovaného výkonu EB môže byť stanovená úroveň napätia a bod pripojenia takto:

- 0.05 až 5MW – vysoké napätie (vn) na úrovni siete, bod pripojenia je realizovaný buď do jestvujúceho vn vedenia, alebo v najnepriaznivejšom prípade priamo do vn prípojnice transformovne 110/22kV,
- 5 až 20 MW – v prípade splnenia technických kritérií bod pripojenia môže byť vn prípojnicou pri najbližšej transformovni 110/22kV. V opačnom prípade sa požaduje bod pripojenia na hladine vvn zaslučkovaním alebo priamym pripojením do vvn prípojnice.
- nad 20MW – napät'ová hladina vvn.

Pripojenie EB do 22 kV siete môže byť vykonané buď zapojením do slučky (T - zapojenie) na existujúce vedenie vn cez vstupnú vn rozvodňu distribuovaného zdroja alebo pripojením do vn prípojnice rozvodne vn pri transformácii 110/22kV [2].

### IV. TEORETICKÉ POZNATKY VÝPOČTU SKRATOVÝCH PRÚDOV V DISTRIBUČNE SIETI

Vzhľadom k tomu, že poruchy v distribučnej sieti spôsobujú nesymetriu v sústave, výpočty poruchových prúdov a napätí sa určujú pomocou symetrických komponentov. Pre tento účel môže byť využitá jedna z najviac používaných metód výpočtu pre rozklad nesymetrickej sústavy na systém symetrických zložiek a to DTF - Diskrétna Fourierova transformácia. Keďže sa jedná o trojfázový systém, je nutné rozložiť nesymetrickú sústavu do troch symetrických zložiek – súslednej, spätnej a netočivej zložky.

Nakoľko v 22kV sústavách prevádzkovaných s neúčinne uzemneným uzlom transformátora vzniká práve pri trojfázovom skrate maximálny skratový prúd, elektrické zariadenia sa dimenzujú na túto hodnotu. V nasledujúcom postupe bude opísaný spôsob výpočtu skratového prúdu a to pre trojfázový skrat, kde je postačujúce určiť súsledné zložky impedancií jednotlivých častí elektrickej siete:

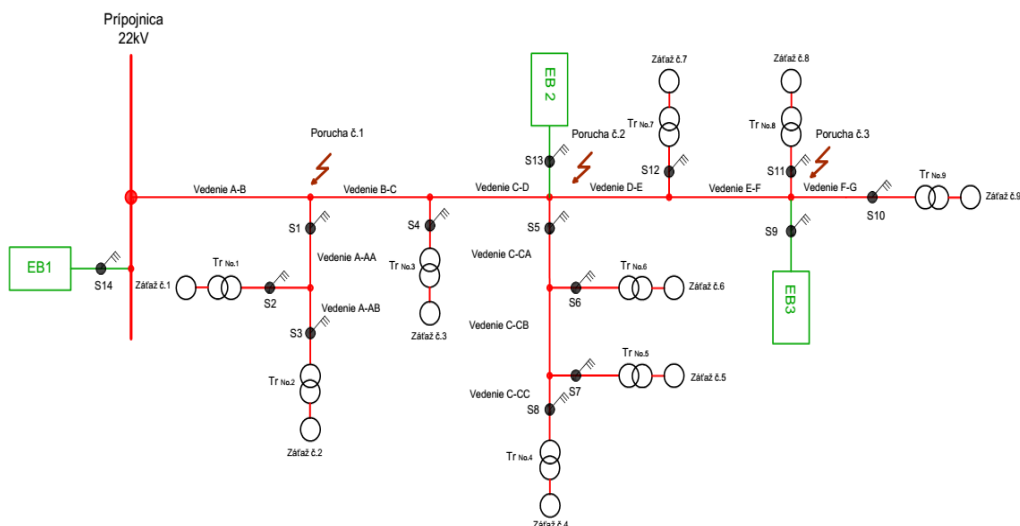
- zostaví sa jednopólová náhradná schéma pre súslednú zložku, kde všetky časti elektrickej siete sú nahradené svojimi impedanciami prepočítanými na vzťahné napätie,
- vypočíta sa výsledná súsledná impedancia, ktorá je daná sériovo-paralelným radením všetkých impedancií medzi miestom skratu a všetkými zdrojmi v sústave,
- výsledná veľkosť trojfázového skratového prúdu je následne určená napätím v mieste skratu a výslednou súslednou impedanciou.

Vyššie boli uvedené stručné teoretické postupy riešenia skratov, detailnejšie postupy a vzorce sú uvedené v [4]. V nasledujúcej kapitole je uvedený modernejší spôsob riešenia skratov a to pomocou výpočtovej techniky, konkrétne v programe Matlab /Simulink [2], [4].

V. VPLYV MIESTA PRIPOJENIA ALTERNATÍVNEHO ZDROJA ENERGIE NA VEĽKOSŤ SKRATOVÉHO PRÚDU V DISTRIBUČNEJ SÚSTAVE

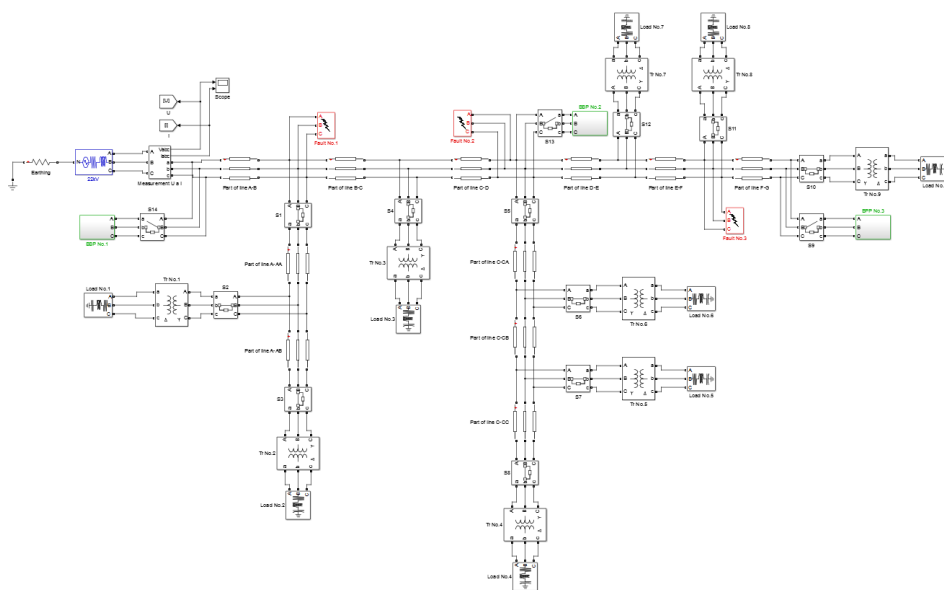
V tejto kapitole sú uvedené výsledky simulácií trojfázové skratu umiestneného na troch rôznych miestach s a bez vplyvu EB pripojenej k sieti. Pre pripojenie EB k sieti boli vybrané tri rôzne pozície, pričom v jednom okamžiku bude pripojená vždy iba jedna EB. Tieto pozície odpovedajú možným miestam pripojenia EB k sieti v reálnych podmienkach. Na základe týchto možných variánt sa zisťovalo, ktoré umiestnenie EB najviac ovplyvňuje veľkosť skratového prúdu v sieti.

Jednotlivé simulácie boli vykonané na uvažovanej časti 22 kV distribučnej siete (Obr.1). Hodnoty transformátorov, vedení a zdrojov boli zadané podľa poskytnutých údajov. Hodnoty jednotlivých odberov sú určené na základe optimálneho zaťaženia v prevádzke. Sústava má deväť odberných miest, každé odberné miesto má svoj spínač na odpojenie, resp. pripojenie.



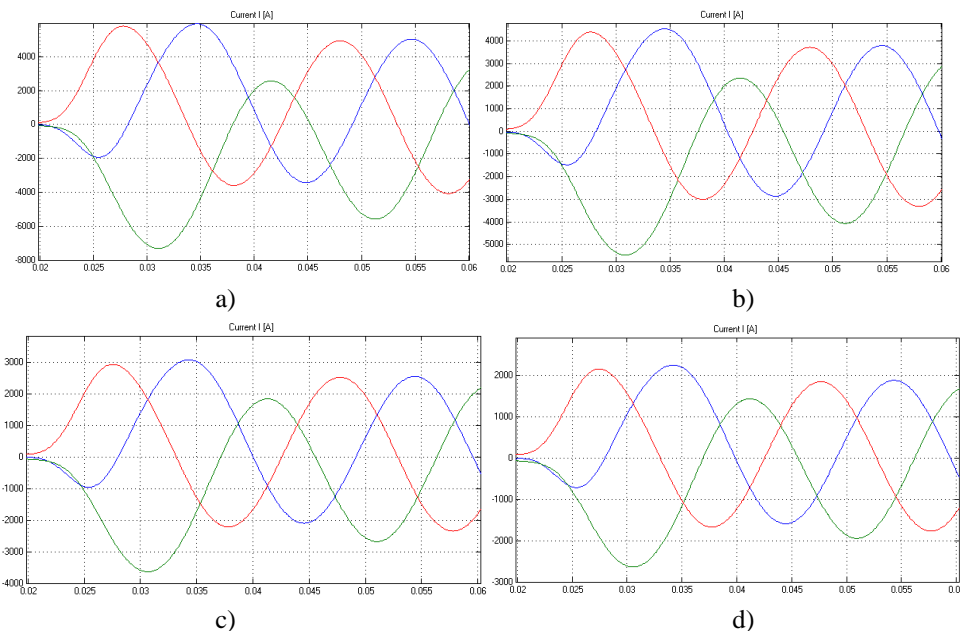
Obr. 1 Uvažovaná časť 22kV distribučnej siete

Na Obr. 2 je znázornený model uvažovanej 22kV časti distribučnej siete vytvorenej v programe Matlab/Simulink. Celá sieť je zložená z vedení, ktorých celková dĺžka je 31,3 km, zatiaľ čo hlavná línia vedenia má 16,6 km. Transformátory v bodoch oderu transformujú napätie z hodnoty 22 kV na 420 V. Hodnoty potrebných veličín, zadané v nastaveniach príslušných modelov prvkov, boli zadané na základe dostupných technických parametrov. V prípade EB ide o zdroj s výkonom 5 MW [3].



Obr. 2 Model uvažovanej časti 22kV distribučnej siete v programe Matlab/Simulink

V prvom prípade bola porucha simulovaná na troch rozličných miestach v sieti bez pripojenia EB. V tomto prípade, veľkosť celkového trojfázového skratového prúdu tečúceho vedením bola rozhodujúca impedancia vedenia zúčastňujúceho sa na skrate, daná predovšetkým jeho dĺžkou. Okrem toho bol simulovaný aj najhorší možný prípad vzniku skratu (bez pripojenia EB). V tomto prípade je najväčšia hodnota trojfázového skratového prúdu  $I''_{K3} = 6.550$  kA. Tento skrat nastal na 22kV prípojnici v elektrickej stanici (Obr.3a), kde následne hodnota skratového prúdu sa postupne znižovala s narastajúcou dĺžkou vedenia (Obr.3b, c, d). Presné hodnoty týchto prúdov sú uvedené v Tab. 1.



Obr. 3 Okamžité hodnoty trojfázového skratového prúdu bez pripojenia EB a) na prípojnici v elektrickej stanici, b) porucha č.1, c) porucha č.2, d) porucha č.3,

V druhom prípade bola porucha simulovaná na troch rozličných miestach v sieti s pripojením iba EB1 do siete. Podľa schémy vyplýva, že po pripojení EB1 na prípojnice rozvodne sa EB1 bude priamo podieľať na zvýšení celkového trojfázového skratového prúdu. Rovnako ako v prvom prípade bude veľkosť skratového prúdu závisieť na vzdialenosti skratu od zdroja (priamo od veľkosti impedancie vedenia, ktorá sa zúčastňuje skratu). Najväčšia hodnota skratového prúdu v tomto prípade bola počas skratu v mieste poruchy č.1 a to  $I''_{K3} = 5.7$  kA, kde príspevok od zdroja EB1 predstavuje 789A. Prúdová hodnota postupne klesala s rastúcou dĺžkou vedenia.

Tab. 1  
 Výsledne hodnoty skratových prúdov pre všetky namodelované situácie

Pripojené zdroje	Miesto poruchy	$I''_{K3}$ [A] (z 22kV)	$I''_{K3}$ [A] (z BPP)	Pripojené zdroje	Miesto poruchy	$I''_{K3}$ [A] (z 22kV)	$I''_{K3}$ [A] (z BPP)
Bez EB (iba 22kV prípojnice)	č.1	5456		EB 2 + 22kV prípoj.	č.1	5448	789
	č.2	3875			č.2	3854	1040
	č.3	2863			č.3	2661	714
EB 1 + 22kV prípoj.	č.1	5262	745	EB 3 + 22kV prípoj.	č.1	5460	850
	č.2	3619	512		č.2	3875	903
	č.3	2624	371		č.3	2861	998

V treťom prípade bola porucha simulovaná na troch rozličných miestach v sieti s pripojením iba EB2. Podľa schémy na Obr. 1 môžeme vidieť, že ak porucha vznikne medzi elektrickou stanicou a miestom pripojenia EB2, poruchový prúd bude tečť z dvoch strán do miesta skratu, v ostatných prípadoch bude skrat napájaný iba z jednej strany. Veľkosť príspevkov skratových prúdov od jednotlivých zdrojov do miesta skratu je priamo závislá od vzdialenosti skratu od týchto zdrojov (tzv. skrat vzniká iba na hlavnej linke vedenia, nie na odbočkách, tým je veľkosť impedancie závislá iba na dĺžke vedenia/impedancie medzi bodom skratu a zdrojom).

V poslednom prípade, porucha bola simulovaná na troch rozličných miestach v sieti s pripojením iba EB3. V tomto prípade môžeme vidieť, že ak je skrat na akomkoľvek mieste (nie odberu, ale na vedení) do miesta skratu bude tečť vždy prúd z dvoch strán.

Presné hodnoty skratových prúdov získaných zo všetkých simulácií sú uvedené v Tab.1.

## VI. ZÁVER

Tento článok rozoberá vplyv miesta pripojenia elektrárne na biomasu spoločne s vplyvom umiestnenia skratu v distribučnej sieti na veľkosť trojfázového skratového prúdu. Pre pripojenie elektrárne k sieti boli vybrané tri miesta pripojenia pričom v jednom okamžiku bola táto elektrárň pripojená iba v jednom mieste. Zo získaných hodnôt simulácií môžeme tvrdiť, že najhorší prípad nastane, ak skrat vznikne v elektrickej stanici na prípojniciach, na ktoré je pripojená aj EB (druhý prípad). Skratový prúd vtedy dosiahol hodnotu cca 7,6 kA. Pri takejto poruche je celá odbočka odpojená od distribučnej siete. Ak by bolo schválené prevádzkovateľom siete napájanie určitého úseku alebo iba vyhradených odberov vyššie spomenutou biomasou, tak posledné dva prípady by sa javili ako najlepšie možné riešenie umiestnenia biomasy v sieti, vzhľadom na zvyšovanie spoľahlivosti prevádzky distribučnej siete.

## ACKNOWLEDGMENT

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied projektom VEGA 1/0388/13.

## REFERENCES

- [1] Biogas power plant design and construction. [online] [cit. 2014-01-05], < <http://www.en-bio.com/biogas-power-plant-design-and-construction> >
- [2] P. Braciník, A. Otčenášová : “The reviewing distributed power sources impact on fault’s localization in 22kV network” Katedra výkonových elektrotechnických systémov, Elektrotechnická fakulta ŽU v Žiline, Advances in Electrical and Electronic Engineering publication
- [3] D. Medved', L. Nemergut: “Modelovanie prechodných javov využitím nástroja Matlab/SimPowerSystem”, Department of Electric Power Engineering at Technical University of Košice, Electrical Engineering and Informatics III publication.
- [4] STN EN 60909-0:Skatové prúdy v trojfázových sústavach, Časť 0: Výpočet prúdov.



# Secured web oriented videoconference system

<sup>1</sup>Roman VÁPENÍK, <sup>2</sup>Katarína FECILAKOVÁ, <sup>3</sup>Dávid CYMBALÁK

<sup>1, 2, 3</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>roman.vapenik@cni.sk, <sup>2</sup>katarina.fecilakova@cni.sk, <sup>3</sup>david.cymbalak@cni.sk

**Abstract** — This paper deals with the means of videoconference systems ran by web browsers. It analyzes certain contemporary videoconference solutions and technologies used for multimedia streaming. It also describes the WebRTC technology terminology and presents the reader with its benefits. It sketches certain options for security enhancements for visual stream transmission via computer network. Based on the proposed concept it was possible to create an experimental web oriented videoconference system with a dual layer security policy in which users are able to communicate via chat and share files with each other besides audio and video communication.

**Keywords** — Security enhancements. Videoconferencing system. WebRTC.

## I. INTRODUCTION

By 2020, 90% of data transmitted over the Internet will have the visual character.

This is confirmed by many studies dealing with this issue. Communication using video is a modern, saving time and effort and it is therefore very popular. Systems dealing with video conferencing are therefore widespread. Each of them has some disadvantages.

Hardware video conferencing systems are the best in terms of image quality and the experience in the communications, but it is often very expensive. Usually it is a device that can be considered as furniture or even to the whole room with equipment according to certain standards. Procure them can't therefore anyone, quite the contrary, is intended close group of people.

Software video conferencing systems are nowadays very popular. Their biggest advantage is price - most of them are in fact free. They are therefore economically available to the general community. Of course, these public versions are not complete and if you are looking for more options of software systems you have to pay extra. However, this addition supplement is incomparably less than the cost of hardware video conferencing systems.

Managing them is challenging. Not everyone knows the complicated architecture of these systems. For some, the problem is no longer the installation itself. In addition, when you install or update may cause problems - the setup file may be infected. The executable file is in fact the best hidden viruses or Trojans.

The aim of this work is to create video conferencing system that is as simple as possible for the end user. Will not require any installation, be easy to use and understand.

## II. GOALS

The main goal of this work was to create a pilot web-based videoconferencing system. The user should be able to share its resources, such as video, audio and data. In addition, the protection of the system should be granted. This task requires several sub-tasks.

The first of these was the analysis streaming technologies. Furthermore, it was necessary to become familiar with existing videoconferencing solutions. Using this analysis a view of the problems of current videoconference systems and their possible solutions are created.

One of the main tasks of this work was to design security extensions, which could be implemented on a web-based videoconferencing system. The result should be a system which allows users to interact with Web browsers. Give them the security of their communications and create a comfortable and easy interface for media sharing.

### III. ANALYSIS

There are many videoconferencing systems providing basic or the advanced options of communication between two or more users. As an illustration in the following chapters are some of them.

#### A. *Software solutions*

##### *Skype*

The most famous software solution of communication between users is Skype. Skype is a communications program allowing running Internet telephony and video. In addition offering the ability to communicate using text messages and it also includes file-sharing. Skype in addition to the basic functions offered for a fee and advanced communication options. The program allows you to make calls to traditional telephone networks or receive calls from mobile networks. After paying the fee even offers the possibility to voice mail. [1]

#### B. *Hardware solutions*

##### *Telepresence*

The best-known hardware solutions include Cisco Telepresence. This is not just an ordinary video conferencing system. Telepresence is a whole set of technologies which allow a person should feel that companion with whom to communicate, is in the same room with him. This system consists of one entire room. The whole room is adapted to standards to be met by all Telepresence room in the world. These include of course layout screens and cameras, microphones, tables and chairs. But in addition includes trivial things such as the material from which are produced chairs, lighting in the room, flower, which should be located in any room or the color and texture of the walls. All these conditions create an amazing atmosphere that truly evokes the impression that communicates who is on the other side of the world, sitting in the same room. [2] The disadvantage of this system is its cost. Device across one room, three large screens, microphones, cameras, etc., is all very expensive and few people can afford it.

##### *Lifesize*

A cheaper alternative of hardware video conference system is LifeSize. He offers a customizable solution that requires a much lower cost than Telepresence. It consists mostly of HD camera, which scans the user and hardware that takes care of the actual processing of video and audio. Then it streams that flow into the network to another user. This hardware can be connected to a television or display device and produces a life-size picture. LifeSize does not offer experience as strong as the Telepresence, because it does not have a special adapted room, furniture or other accessories, but is a good compromise because of its cost and quality of services. [3]

#### C. *Web-oriented solutions*

##### *Hangout*

Addition to typical software and hardware solutions, there are largely extended web-based video conferencing system. One of them is the Hangout from Google. Hangout allows communication of two or more users. The service is accessible from the websites of Google mail – Gmail, through websites Google+ (Google's social networking site) or through mobile applications on operating systems Android and iOS. To establish a connection user must be logged in your account, which must give Google a few of own personal data. This service is linked with social network Google+, which is recorded all communications made through the Hangout. Google Hangout uses its own proprietary protocol, not the standard protocol XMPP. So Third party products can't establish a connection with Hangout. [4]

#### D. *WebRTC*

WebRTC is an open project which allows you to communicate in real time via a web browser, using a simple JavaScript. This project is in no small measure supported by three web browsers - Google Chrome, Mozilla Firefox and Opera. [5]

WebRTC is a videoconferencing system that is based on communication between two web browsers. Communication allows two or more users. It does not require installation of additional plug-ins or any additional applications. Provide the communication using video and audio, through text messaging and user's desktop sharing, and part of the system allows sharing files between users. [6]

The biggest advantage of this technology is its simplicity. With a few lines of source code written in JavaScript can be used WebRTC functionality. However, the use WebRTC API requires initial configuration as a signaling, and search for a client, security settings, and many others. In addition, for full functionality of WebRTC is necessary some new layer protocols. However, most of the work carried out by the browser itself, which processes data in three areas using three basic components of WebRTC:

- `MediaStream` - accessing the local input devices
- `RTCPeerConnection` - setting the remote audio and visual data,
- `RTCDataChannel` - sharing data between end points. [6]

WebRTC architecture is made up of a number of standards, including the browser and application API, as well as many other protocols and data formats required for correct functionality. For this, it was desirable to develop and introduce new rules to be implemented in web browsers. [5]

#### IV. SOLUTION

##### A. Architecture of system

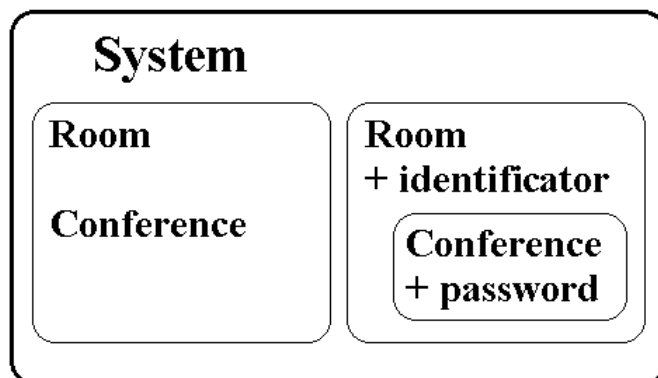


Fig. 1 Architecture of system

Fig. 1 shows the system architecture. As is visible, it consists of two parts - the public (left panel) and private (right). The public part that has no security consists of one common room and conferences which are visible to each user of the system. Private contains two layers, each with its own security.

Each private room has its own unique identifier - a token, which is clearly described. When room is created, this identifier is generated using pseudo-random function. Without it, the entrance to the room excluded because URL address of room contains this identifier. The second layer of the private part of system is the conference. It is protected by a password that is requested from the user who creates a private conference room. This password is, of course, required by users who want to join the conference. The client who wants to access the private conference needs from the founder or already connected users get a unique token (or URL) of private room and password.

##### B. Types of conferences

In this system, there are four types of conferences: many-to-many, one-to-many, one-to-one and one-way. They differ by participant's interconnection and direction of distribution of signal between them.

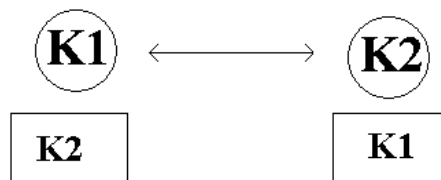


Fig. 2 One-to-one conference type

The first type is a one-to-one (Fig. 2). It's a classic case of a videoconference system for two people. Stream flows from the client K1 to the second and from the client K2 to the first. Both therefore shows stream by on the other end connections.

Conference type "one-to-many" (Fig. 3) is applicable in relation teacher-students or instructor-students. K1 client as a "teacher" sees all other clients (students), so from all of them stream is delivered to him. But "students" does not see each other. The only stream that is transmitted to them is from the main client K1.

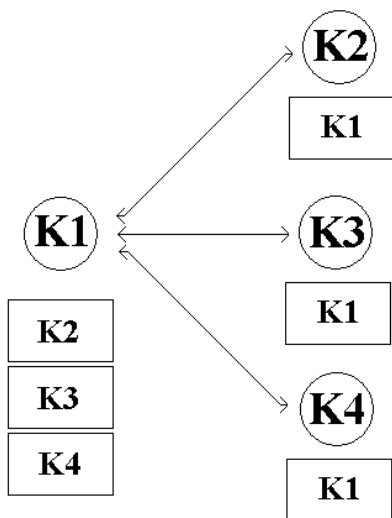


Fig. 3 One-to-many conference type

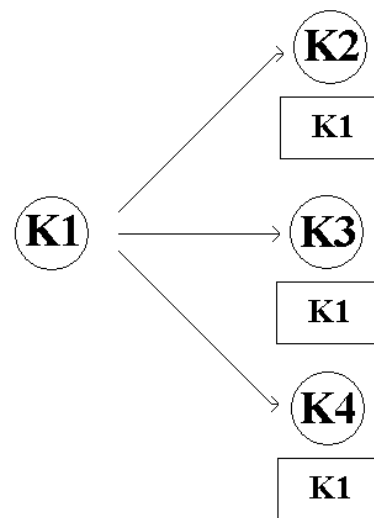


Fig. 4 One-way conference type

One-way conference (Fig. 4) is like typical television broadcasting. Client K1 sends its source to all connected participants. However, he does not receive any stream from the users. At the one-way conference, own resources displayed to the user when the first client is connecting to follow the stream. Receiving participant does not have his own stream displayed at all. It is therefore just like watching TV, when we see images transmitted from the studio, but not own image. And people in the studio do not see the picture of our living room.

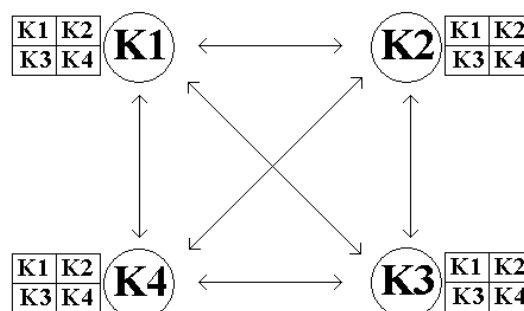


Fig. 5 Many-to-many conference type with Full Mesh topology

The last type of conference is many-to-many, in which all clients can see streams from all other clients. In this type there are two solutions to the distribution of information. One of them is called "Full Mesh" topology (Fig. 5), where information is distributed from each client to each. In this case, there is a problem with the procedural load on all participants. Each of them must process and render stream from all other clients as well their data must be sent to all other machines. This may be a problem in case of greater number of users and a weaker computing machine.

The second solution is topology Star (Fig. 6). Out of all participants choose one that has the best computing resources and become supernode. All streams from users are redirected to this client and they will distribute to other client. Other clients are not so loaded, because they will send and receive streams only from the supernode.

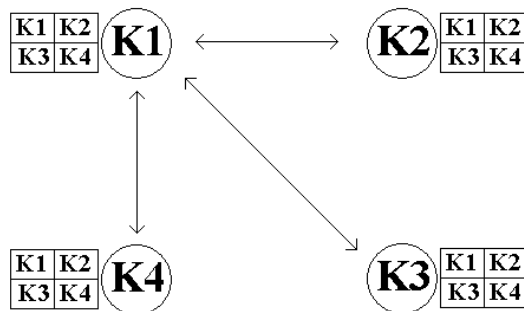


Fig. 6 Many-to-many conference type with Star topology

## V. CONCLUSION

The task of this work was to create a system that would allow communication between amateurs and professionals. This system was built as a web videoconference system with security features. It is used to allow communication in four forms - video and audio form, in the form of text chats and files sharing between users.

An important feature of this system is its security. In addition to the primary security control of information exchange and of course the secure transmission of video, audio and data, the system was extended to the two-level security videoconferencing virtual rooms.

When you create a private room is used a unique token that uniquely describes a given room. The user who wants to join in this room must know this token, respectively founder of this virtual room have to give it. In the room may be several conferences, which are differentiated from each other by their names.

When accessing to a particular conference user is prompted for the conference password. This password is the second level of system security. User's password must match those that were specified when creating the conference.

The videoconferencing system has been tested on most operating systems - Windows 7, Windows 8, Unix and Android. Thus, it is multiplatform and in addition is functional on multiple distribution of web browsers.

## ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

## REFERENCES

- [1] 1. BASET, Salman A.; SCHULZRINNE, Henning. An analysis of the skype peer-to-peer internet telephony protocol. arXiv preprint cs/0412017, 2004.
- [2] 2. SZIGETI, Tim, et al. Cisco TelePresence Fundamentals. Cisco Press, 2009.
- [3] 3. PAN, Hui. Telecom Mergers & Acquisitions Monthly Newsletter November 2009. Information Gatekeepers Inc.
- [4] 4. XU, Yang, et al. Video telephony for end-consumers: measurement study of Google+, iChat, and Skype. ACM, 2012. p. 371-384.
- [5] 5. JOHNSTON, Alan B.; BURNETT, Daniel C. WebRTC: APIs and RTCWEB Protocols of the HTML5 Real-Time Web. Digital Codex LLC, 2012.
- [6] 6. GRIGORIK, Ilya; Secure Communication in WebRTC. High-Performance Browser Networking. O'Reilly Media, 2013

# 8-bit ATMEL microcontroller based Street view bicycle

*Tomáš ECKER, Peter FECILAK, Katarína FECILAKOVÁ*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

tomas.ecker@student.tuke.sk, peter.fecilak@cni.sk, katarina.fecilakova@cni.sk

**Abstract** — The aim of this paper is to design and implement structural movement device on Google Street View. For this purpose is used ATtiny2313 microcontroller, mounted in a housing on the board with a USB cable and connectors for connecting sensors. The sensors are placed on the bike and allow generating signals for moving forward, left and right during cycling. Construction of this device allows movement in Street View application through cycling. The whole system is implemented as a HID device, thereby is obtained an expanded Street View of the new peripheral device.

**Keywords** — Google Street View, HID device, ATtiny2313, bicycle, sensors for generating signals

## I. INTRODUCTION

Nowadays we are surrounded by a different of technical convenience and software solutions, which largely make our life easier. A very useful application is Street View, which was designed and implemented to its present by Google. Possibility of this service has become an inspiration for this work. What if the person could move around the streets other than using a keyboard and mouse? This led to the creation of Street View bicycle.

Brain of most devices is the microprocessor that performs the instructions and process information which are important for performing certain tasks. There are many types and brands, but for the purpose of this work is used AVR microcontroller ATtiny2313. This microcontroller belongs to the 8-bit chips with small measurements and sufficient parameters to achieve the objectives.

Street View uses peripheral devices such as mouse and keyboard to move around different places. Nevertheless, it is possible to use other solutions, which definitely includes the Street View bike. This system allows extend peripheral devices by the new HID device. Human interface devices are currently one of the most used devices that are connected to the USB port of a computer and allow full use of the various services and applications. Street View bicycle system allows the generation of events for Google Street View during cycling. Thanks to this solution is movement in this service more entertaining, and it used physical activity of person. This device in today's high-technology world, when people become more comfortable and lazier, could mean a leak from this stereotype and could lead them to a minimum level of physical activity. With this activity they can combine work with pleasure which in addition to this relatively light physical activity which uses Street View may acquaint with interesting historical, cultural and natural attractions.

## II. STREET VIEW SERVICE

Street View was developed by Google to capture and mapping streets and different objects and places in worldwide. This service was first launched May 25, 2007 and was gradually available in other countries in worldwide. Nowadays it is available on every seven continents. How it is explained on Google Maps web page [5], individual places are captured as 360 degrees panoramic shots, which are made with several cameras. These shots are overlapped to

prevent spaces in the resulting panoramic shot. For scanning Street View uses, as Google Maps says [4] :

- cars
- snowmobiles
- trucks
- tricycles
- trekkers

#### *Mouse controller*

With mouse, person can click on place where he likes to go. The more places are captured, the more options are offered for user and he is able to go to the places, where he normally can't.

#### *Keyboard controller*

Keyboard using allows more fulfilling for control. For me it is better solution as mouse controller. Someone can imagine, that he playing a computer game.

### III. AVR ARCHITECTURE

These types of microprocessors are fast and performing. AVR microprocessors are based on RISC [2] architecture (Reduced Instruction Set Computer). It means that their instruction set contains only basic simple instruction unlike CISC (Complex Instruction Set Computer) [3]. They have separated data and program memory. Data memory is SRAM (Static Random Access Memory) or EEPROM (Electrically Erasable Programmable Read-Only Memory) type and program memory is FLASH type.

#### *ATtiny2313*

This microprocessor belongs to the ATtiny ones, which produced Atmel. It is low-voltage, eight bit AVR microprocessor based on RISC architecture. ATtiny2313 has these specifications, which are shown in free data sheet [1]:

- 2K bytes programmable memory
- 128 bytes EEPROM memory
- 128 bytes SRAM memory and more other

It has also 20 pins. Two of them are GND and VCC and remaining ones belongs to ports. ATtiny2313 has three ports: PORTA, PORTB and PORTD.

### IV. CONSTRUCTION OF DEVICE FOR MOUNTING MICROPROCESSOR

The main part of whole system is little board with ATtiny2313 microprocessor. This microprocessor is mounted in case on this board. This solution is appropriate because of picking out and connecting to the programmer, but also for breakdown. On the board are placed three keys, which we can use like keyboard in case that we don't use sensors from bicycle. It also has connectors to connect sensors from bicycle. They are connected to the pins of microprocessor.

- one connector is connected to the PB0 pin of the microprocessor
- one connector is connected to the PB1 pin of the microprocessor
- one connector is connected to the PB2 pin of the microprocessor

To connect this board to a computer, on the board is placed USB cable. Four wires are different color and connected to the pins follows:

- black wire to the GND pin
- red wire to the VCC pin
- green wire to the PD2 pin
- white to the PD5 pin

Whole device is shown on the picture Fig. 1.

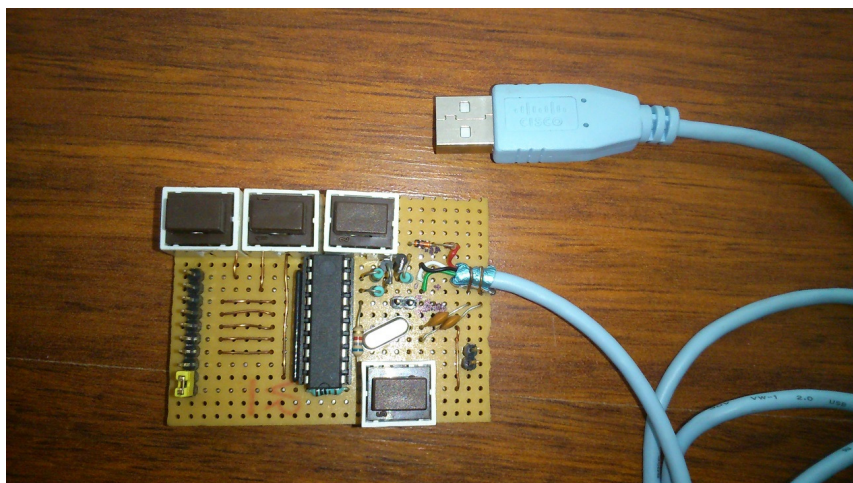


Fig. 1 Device with microprocessor and USB cable

## V. BICYCLE AND LOCATION SENSORS

On the bicycle there are three sensors for generating signals for Street View application. The figure Fig. 2 represents sensors to detect steering wheel movement. There are used two conductive plates, every placed on one side of steering wheel. On the construction of bicycle is one plate, which is used for closure and generating signals to move left or right in Street View service.



Fig. 2 Sensors on steering wheel

To generating signal for forward moving is used sensor from cycle computer. This sensor is placed on back frame next to the wheel and one beam of wheel has magnet as shown on the figure Fig. 3. When magnet and sensor are overlaid, signal for forward move is generated.



Fig. 3 Sensor on back frame and magnet

Each of these three sensors has connectors, which allow connect them to the board and subsequently moving in Street View during cycling.

## VI. SOFTWARE FOR OPERATING THE MICROPROCESSOR

The whole program is written in C. One of the most important file is `usbconfig.h`. It contains symbolic constants for identifying USB devices. In this file are two important constants, which are `USB_CFG_VENDOR_ID` and `USB_CFG_DEVICE_ID`. These two values represents vendor ID and device ID and on this basis is our device recognizes as HID device.



Computer then install needs drivers and device is ready to use. The program contains three constants `LEFT_ARROW`, `RIGHT_ARROW`, `UP_ARROW`, which represents three keys on standard keyboard, and they are in two-dimensional array `keyReport[]`.

Every pin has own index and one of this constant is assign to the one pin of microprocessor. Index number is assign to every pin in `keyPressed()` function, if any key or sensor is pressed. In this way, is provided, that after sending signal from one of the sensors is generated key code, and we can move in the Street View application by cycling.

## VII. CONCLUSION

Topic 8-bit ATMEL microcontroller based Street view bicycle required construction of the device for recording impulses and their transfer via USB and implementation of software solution. These steps were fulfill through the completed board with embedded microprocessor ATtiny2313 and brought out the USB cable to connect it to the USB port of a computer as a HID device. To the board can be connected to sensors that allow generating impulses that are processed and properly used through the software solution. Device is created for design - programming project. It is not just certain application programming, web interface or another software solution, but to create the hardware on which it is then possible implementation. The theme of this work provides a comprehensive solution to the problem in information technology. The actual hardware with embedded microprocessor allows use in other issues. For example device to automatically fill the field for the password in the various systems, that relieve a person from remembering a large number of passwords.

The whole Street View bike system is designed to move in Google Street View during cycling, on which it was constructed. This system can be expanded in the future to ability record cycling speed and then simulate faster, or slower move in Street View application. The system was designed and implemented as a HID device, which has obtain to expansion of existing devices such as mice, keyboards, game controllers. This solution represents the input peripherals that provide movement in Street View. The fact that the designed system is implemented as a HID device allows you to extend the supported peripherals for this service and provides a new way of moving through the streets and places that this service provides.

## ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] ATMEL: 8-bit AVR Microcontroller with 2K Bytes In-System Programmable Flash [online]. 2010. [Cited. December 14, 2013. <http://www.atmel.com/images/doc2543.pdf>]
- [2] DŘINEK, Milan. RISC verus CISC architektúra [online]. 2000-01-12. [Cited. November 26, 2013]. [http://avr.hw.cz/architektura/risc\\_cisc.html/](http://avr.hw.cz/architektura/risc_cisc.html/).
- [3] Části jednočipového mikro počítače. [Cited. November 26, 2013]. <http://www.programovani.howto.cz/casti-jednocipoveho-mikropocitace/>.
- [4] Google Maps: Behind The Scenes Street View. [Cited. October 27, 2013]. <http://www.google.com/intl/en/maps/about/behind-the-scenes/streetview/>.
- [5] Google Maps: Frequently Asked Questions. [Cited. October 25, 2013]. <http://www.google.com/help/maps/mapcontent/streetview/faq.html#q4>.

# Development of automated methods for network traffic filtering depending on a client station behavior

*Barbora CHVÁLOVÁ, Peter FECILAK, Katarína FECILAKOVÁ*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

barbora.chvalova@tuke.sk, peter.fecilak@cni.sk, katarina.fecilakova@cni.sk

**Abstract** — The aim of this paper is to develop a method for monitoring and filtering the traffic which is generated by a client station. The focus is on the automation of this process after having chosen suitable technologies. Selection of technologies applies to areas such as packet capturing, remote management of network devices and filtering of the packet flow. As a form of implementation, an application with a graphical user interface is programmed for better handling and usability. The final application is possible to be utilized in smaller companies that have deployed new network infrastructure and need to acquire inner security policy. This application helps to ensure basic host monitoring in a network and define the standard network traffic.

**Keywords** — filtering, client traffic, anomaly, intrusion detection, remote management

## I. INTRODUCTION

Nowadays, the client station behavior is not a starting point for creating network security policy. Every company has its policy firmly set and monitors traffic based on those rules. Let us suggest that a new company exists without such policies yet defined, with no firewall rules. In this case, policy for network traffic is “permit all” by default and therefore deny statements will be created in the real time.

This work answers the question of creating boundary between standard and anomaly traffic. The work focuses on handling IP traffic and remote management of network devices that are responsible for filtering packets. This process shall be automated in order to achieve high utility and effectiveness.

## II. GOALS

The goal of this work is to analyze and choose between methods of capturing packet for the purpose of learning standard traffic in the first place and monitoring anomaly traffic in the second place. Automation of this process is necessary and will be provided by an application with graphical user interface.

## III. ANALYSIS

In order to complete the analyzing task, options of fields of packet capturing, remote management and anomaly detection need to be taken into consideration. The final method consisting of these technologies will lead to application of filtering traffic. The key device here is a server which has to receive and handle packets from client stations, learn that traffic and after that process it. It also generates ACL configuration and copies it to Cisco device.

The first problem addressed represents capturing of communication from the client station (see Figure 1 part 1a and 1b). There are a few possibilities how the server could receive copies of the clients traffic. Firstly, it is a packet capture library PCAP [1]. It could be used in a socket communication between a client and the server. The client runs a daemon that makes copies of the client's traffic and sends it to the server. This option might cause exceeded traffic and losing of packets. The PCAP library can also be used as a remote packet capture [2] from the server. Second solution might be a hardware packet capturing. It is represented by TAP network devices, port aggregators, as well as network capture cards. Finally, regarding the network

devices EPC (Embedded Packet capture) and SPAN (Switched Port Analyzer) should be considered. EPC [3] is a Cisco IOS feature for routers and is able to capture packet to pcap file directly on a device. The file is then exported for analysis to the server. SPAN [4] (also port mirroring) could copy traffic from source ports to one destination port by mirroring traffic. Server connected to the destination port receives copy of all traffic of those source ports.

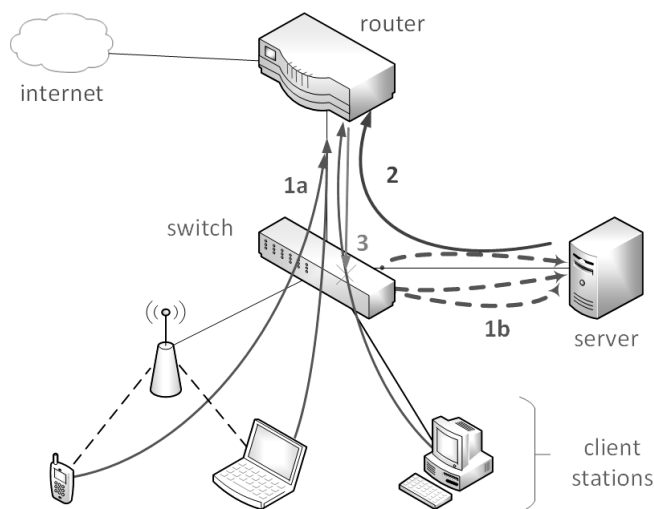


Fig. 1. Example of network topology

The second problem would be an analysis of client's traffic per packet. The PCAP library is able to handle every received packet either in online stream or offline pcap file. Header of packet contains all needed information such as a source and destination IP address, transport protocol and port. All such information is needed to look for anomalies. In case of analyzing payload – data of packet, deep packet capturing [5] and analysis should be used. The header information could be used to filter packet by using Access Control Lists (ACL) [6] – see Figure 1 part 3.

To approach the final problem we need to decide how to remotely configure network device within application (Figure 1 – part 2). Since this process should be automatic, such a method has to be effective. There are two technologies at hand: Secure Shell (SSH) and Simple Network Management Protocol (SNMP). SSH is a secure version of telnet protocol and it is basically a connection via virtual terminal to a management IP address of a device. In order to achieve a secure administration, this option is preferred. SNMP [7] is a protocol based on requesting and setting specific information about devices where information is stored as objects within hierarchical information tree called MIB (Management Information Base). Such a tool could be used for copying configuration from device to device [8], what could be used on a larger scale than SSH.

#### IV. SOLUTION AND RESULTS

The solution is complex in a way that besides final application, there is a need to configure network device. The configuration of network device has to be complementing the application settings. The implemented application aFoNT (automated Filtering of Network Traffic) captures packets that are mirrored to destination SPAN port connected to the server. The application then handles each packet according to the chosen mode. The modes are selected by the user of application and it could be either mode called LEARNING or MONITORING.

Learning mode means that all captured packets are dumped to the file until the user stops the capturing. Afterwards, the file is parsed and only pairs of source – destination IP addresses are written to a new file. It simplifies the output for purposes of application because only IP flows are now needed for monitoring.

Monitoring mode uses the parsed file and compares real time traffic against it. The application written in Java programming language is able to keep up to handling and analyzing packet load. In case of an anomaly, a new programming thread is started to create ACL configuration and send a configuration file to be merged with a switch running the configuration. For that the SNMP protocol is used.

The aFoNT application is working as a kind of Intrusion Detection System (IDS). IDS is a powerful security feature which is able to recognize any intrusion and anomaly and either inform the administrator of the network or block the detected intrusion.

#### A. Software and Network features

The main network features are SPAN and SNMP. SPAN feature, as described in the Analysis, is able to mirror all/partial traffic according to network settings. The traffic mirrored from source ports could be limited to either received (RX) or transmitted (TX). Destination port connected to the server has no other traffic and no IP address. Therefore the server is not able to communicate through this interface. As seen in the Figure 2, at least one other network interface is needed in order to manage the switch.

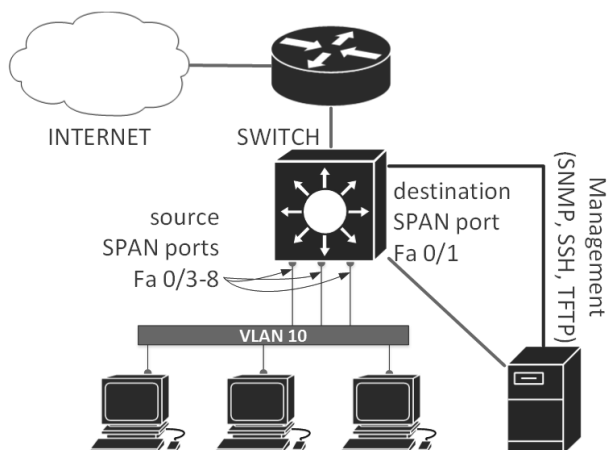


Fig. 2. Testing topology

SNMP forms a communication canal between a network device and the application server. SNMP works with TFTP transfer protocol for file transfer and uses objects in MIB tree called CISCO-CONFIG-COPY-MIB with OID 1.3.6.1.4.1.9.9.96.1.1.1.1 and type *ccCopyEntry* [8].

```
CcCopyEntry ::= SEQUENCE {
    ccCopyProtocol          ConfigCopyProtocol,
    ccCopySourceFileType   ConfigFileType,
    ccCopyDestFileType     ConfigFileType,
    ccCopyServerAddress    IPAddress,
    ccCopyFileName         DisplayString,
    ccCopyEntryRowStatus   RowStatus,
}
```

SNMP management station sends objects in a sequence such as `snmp setRequest`. Instances of these objects are initialized and then the whole table entry is activated by the last object. Then the configuration file is downloaded to the network device and merged with running configuration. In this moment, the updated ACL blocks the traffic defined by source-destination IP pair.

As a software feature goes, implementation of the final application is based on Java programming language because it is a multi-platform and is utilized for network functions. Other java software open-source libraries are used to complete other necessary tasks. Library `jNetPcap` [9] wraps functions of `libpcap` and `wincap`. It uses `pcap` native methods and on the top of that it adds its own methods to handle packets. `jNetPcap` methods are used to capture traffic per packets and analyze its header.

Other open-source library is used for SNMP – `SNMP4J` [10]. The simple SNMP manager was created and works together with a SNMP agent on the network device. Community strings are defined to secure access to the device. SNMP uses TFTP for file transfer. A part of the application is embedded TFTP server based on Verinec project [11] in Java.

#### B. Application's results test

Simple multiple-ping test shows a breaking point when the application evaluates an IP

address as an anomaly and apply ACL on the switch. The output shows:

```
C:>ping cisco.com -n 100
Pinging cisco.com [72.163.4.161] with 32 bytes of data:
<output omitted>
Reply from 72.163.4.161: bytes=32 time=154ms TTL=231
Reply from 192.168.2.1: Destination net unreachable.
```

As a result, we can see a ping reply from VLAN 10 gateway, where the ACL were applied and traffic was stopped. VLAN 10 is also a client's network (see Figure 2). The applied ACL could be seen in a running configuration of managed switch. The output also shows matched packets in the moment of show command:

```
Switch# show ip access-lists
Extended IP access list afont1
1 deny ip host 192.168.2.102 50.31.164.0 0.0.0.255 (19 matches)
2 deny ip host 192.168.2.102 23.62.237.0 0.0.0.255 (429 matches)
3 deny ip host 192.168.2.102 72.163.4.0 0.0.0.255 (63 matches)
1000000 permit ip any any log (6195 matches)
```

A sequence of all steps, from learning phase to traffic monitoring and recognizing anomalies, is possible to track in application's GUI (see Figure 3).

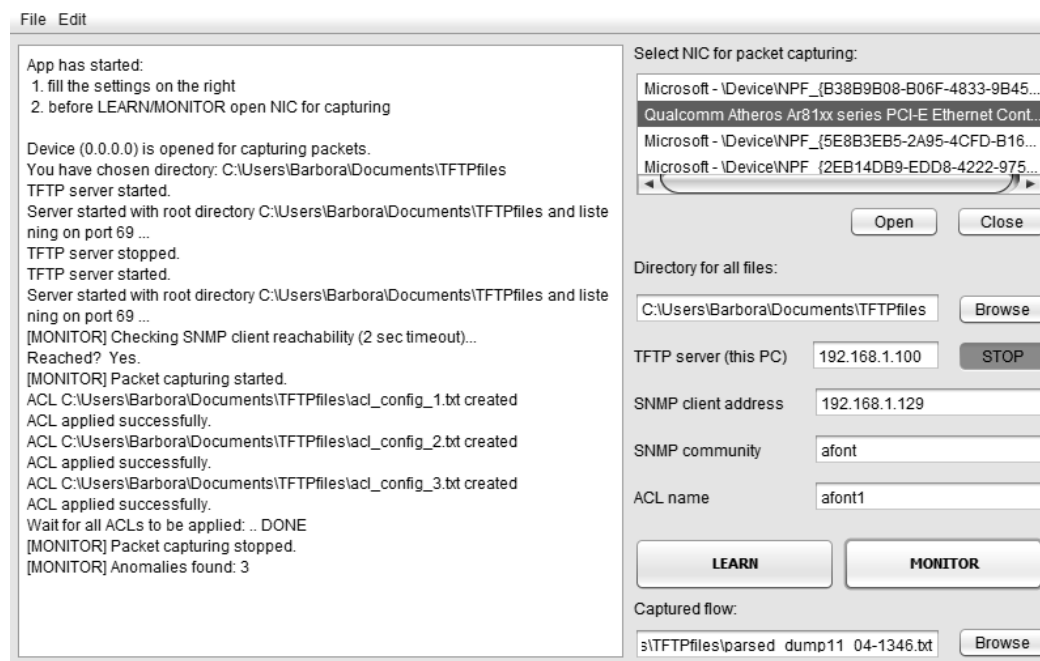


Fig. 3. GUI of application aFoNT with simple test of functions

## V. CONCLUSION

The aim of this work was to define and handle anomaly network traffic generated from the client station. After analyzing options for monitoring and filtering network communication and possibilities of remote management of Cisco devices, I have created software application that handles automated filtering of network flow by comparison of learned and real-time network traffic. If there is any anomaly found, the ACL list is remotely configured and anomalies are blocked.

Traffic handling is based on a packet capture library called "pcap". Its OS forms - libpcap for Linux and winpcap for Windows, are wrapped in jNetPcap Java library. Other possibilities for packet capture are represented by EPC or hardware packet capturing. To manage the device remotely, I used SNMP remote configuration based on copying files to a running configuration of the network device.

The prototype application is working in a testing environment. To develop an application

for production network and a real company use, certain changes, smaller or bigger, are necessary. Firstly, I recommend using special hardware network cards for detailed packet capturing. Secondly, because SNMP slow processing represents this method's limitation, there is a need to improve the remote configuration either by improving the SNMP or implementing a SSH protocol.

#### ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

#### REFERENCES

- [1] WINPCAP: *WinPcap*. [2013-04-25] Internet source: <<http://www.winpcap.org/>>
- [2] WINPCAP: *Remote Capture*. [2013-06-25] Internet source: <[http://www.winpcap.org/docs/docs\\_40\\_2/html/group\\_\\_remote.html](http://www.winpcap.org/docs/docs_40_2/html/group__remote.html)>
- [3] SEMPERBONI, Fabio: *EPC - an Embedded Packet Capture*, USA, 22.06.2011. [2013-12-10] Internet source: <<http://www.ciscozine.com/2011/06/22/epc-an-embedded-packet-capture/>>
- [4] CISCO: *Catalyst Switched Port Analyzer (SPAN) Configuration Example*. [2013-11-02] Internet source: <[http://www.cisco.com/en/US/products/hw/switches/ps708/products\\_tech\\_note09186a008015c612.shtml](http://www.cisco.com/en/US/products/hw/switches/ps708/products_tech_note09186a008015c612.shtml)>
- [5] *Deep packet capture*. [2013-04-25] Internet source: <<http://www.techopedia.com/definition/27459/deep-packet-capture-dpc>>
- [6] CISCO: *Access Lists Introduction* [online]. USA [2013-12-20] Internet source: <[http://www.cisco.com/en/US/tech/tk648/tk361/tk821/tsd\\_technology\\_support\\_sub-protocol\\_home.html](http://www.cisco.com/en/US/tech/tk648/tk361/tk821/tsd_technology_support_sub-protocol_home.html)>
- [7] *SNMP Fundamentals*, USA. [2013-12-23] Internet source: <[http://www.cisco.com/en/US/docs/switches/wan/mgx/ses/software/ses\\_pnni\\_r3.0/configuration/guide/oricgape.html#wp1003191](http://www.cisco.com/en/US/docs/switches/wan/mgx/ses/software/ses_pnni_r3.0/configuration/guide/oricgape.html#wp1003191)>
- [8] *CISCO-CONFIG-COPY*. [2014-3-22] Internet source: <<ftp://ftp.cisco.com/pub/mibs/v2/CISCO-CONFIG-COPY-MIB.my>>
- [9] SLY Technologies: *jNetPcap Home*. [2014-1-8] Internet source: <<http://jnetpcap.com/>>
- [10] *SnmP4j Home*. [2014-3-22] Internet source: <<http://www.snmp4j.org/>>
- [11] Prof. ULTES-NITSCHKE, Ulrich - JUNGO, Dominik - BUCHMANN, David: *Verinec project*. [2014-3-22] Internet source: <<http://verinec.sourceforge.net/>>G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

# Environment for creation and reviewing of learning materials in the form of LTI application for Canvas system

*Monika MITRIKOVÁ, Katarína FECILAKOVÁ*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

monika.mitrikova@student.tuke.sk, katarina.fecilakova@cni.sk

**Abstract** — This paper presents the issue of implementation of an LTI application and its integration into the Canvas system. The first part is an analysis of the technologies used for implementation. For the needs of presenting learning content as open source wiki software MediaWiki was chosen. This software was deployed on a school server and extended by software package LTI4MediaWiki, which allows MediaWiki to become an LTI Tool Provider. This software is tested and prepared for use in Canvas system like environment for creation and reviewing of learning materials.

**Keywords** — LTI application, Canvas, Wiki, Learning Management System

## I. INTRODUCTION

Education is a form of learning, where some important skills and knowledge is transferred from one generation to another by teaching or training. In these days, people all over the world use web to make the educational process more efficient and available to everyone. Developers all over the world are working on breathtaking learning applications, platforms and many software tools which can help to understand subjects better, because that is main goal of the educational process.

The main reason for picking this thesis was the opportunity to gain experience in web development, network services and help all networking academy students get better learning experience.

## II. USED SOFTWARE TECHNOLOGIES

### A. CANVAS by Instructure

Canvas is a new, open-source Learning Management System (LMS) developed by Instructure Inc. in 2011. The main idea of this company is focusing on user itself and making the experience from using their product effortless. Therefore, Canvas is concepted on user-friendly way and its interface is intuitive and easy to use. [1]

Everybody who wants to develop for Canvas will come through with REST API (Representational State Transfer Application Program Interface). The REST API is all about reading and/or writing Data Stored in Canvas. This functionality allows developers to work with data by four HTTP methods GET, POST, PUT and DELETE to execute different operations. Canvas allows developers to extend functionalities by adding external tools which can be configured on a course or account level, and can be added to course modules or used to create custom assignments. In the past, a big problem for developers was that one external application could have been used only in one LMS, for which it was created. This problem was solved by IMS Global Learning Consortium by developing a specification called Learning Tools Interoperability (LTI).[2]

### B. LTI Standard

LTI stands for Learning Tools Interoperability, which is a specification developed by IMS Global Learning Consortium. The key concept of LTI is to establish a standard way of integrating rich learning applications, often remotely hosted and provided by third-party services, with platforms like LMS. In LTI these learning applications are called Tools, developed by *Tool Providers* and placed into *Tool Consumers*. [3]

The main thing for developing LTI application is the identity assertion which is mostly used to allow the user to be automatically signed in and directed to specific context when application is rendered. Tools are launched inside the learning platform in an iframe, so user feels like they are a native part of the system. The consumer and provider of tool have some predefined relationship via a *consumer key* and *shared secret* which provide a secure way to verify messages passed between systems. All messages are signed with an *OAuth signature*, which is a standard way to share personal information with third-party services without sharing user's password. Identity assertion happens through an HTTP POST request from the Learning platform to the Learning tool (see Fig.1).

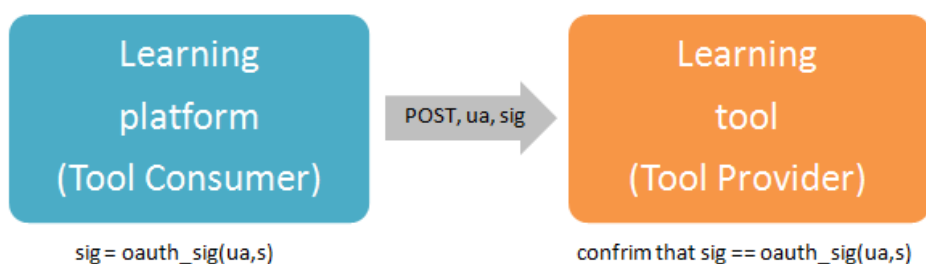


Fig. 1 Identity assertion coming from the learning platform (Tool Consumer) to the learning tool (Tool Provider). OAuth signature contains two parameters, user attributes (ua) and shared secret (s).

This POST request has to happen in the user's browser after launching by submitting a form. Some required and some optional parameters are sent as a part of the POST request. [4][5]

### C. Open source wiki software

For the needs of the creation and review environment of learning materials wiki software was picked. A wiki is a web application which allows people to add, modify, or delete content. This content is usually written using a simplified markup language or rich-text editor. Context of wiki is managed in a collaborative way, so everybody who has the rights (it depends on privacy settings of each wiki) can add, or edit content.

As a conclusion from analysis of the most widely used open source wiki environments, the software package MediaWiki was chosen. This free software wiki package written in PHP was originally used on Wikipedia- most famous free internet encyclopedia. MediaWiki is a free server-based software licensed under GNU General Public License (GPL). It is a powerful and scalable software which uses PHP to process and display data stored in a database. When user submits an edit to a page, it is written to the database, but without rewriting the previous versions on the page. This feature is used for preventing any unwanted content to be displayed on the page. MediaWiki can manage image and multimedia files, and also supports caching. [6]

## III. IMPLEMENTATION

### A. Installation and configuration

Before any installation, we need to set up our server according to MediaWiki prerequisites. First of all, we need to have a web server, such as Apache or IIS to serve the requested pages to the client browser. PHP in version 5.2.4 or later is required to run the software. To store pages



and site data, a database server is required, such as MySQL 5.0.2 or later. To meet these and some additional requirements, our school server was set up with following software:

- Ubuntu version 5.27
- PHP version 5.2.4
- Apache version 2.4.7
- MySQL version 5.0.51a

Installation of MediaWiki was launched by the command prompt `sudo apt-get install mediawiki` for an automatic download and installation of MediaWiki software package.

After successful installation we need to run configuration form in web browser by typing in our domain address of mediawiki package. This configuration will generate LocalSettings.php file which contains all of the main settings of our wiki. For our needs, we have chosen name Netpedia for our wiki.

What we do next, is extending our Netpedia with LTI standard. For this purpose we use free open source software package LTI4MediaWiki. This plug-in allows the MediaWiki to become an IMS Basic LTI tool provider. This software was developed by a group called Learning Apps. After placing this software into our mediawiki extensions folder, there is some additional authentication extension scripts which need to be inserted into LocalSettings.php folder.

Last step is changing *shared secret* and *consumer key* in attached configuration file. These two variables will be used for configuration Netpedia in system Canvas. Everybody who wants to add Netpedia as course module will need to past them into Canvas configuration form.

### B. Netpedia as external tool in Canvas

After logging in the Canvas, External Apps can be configured on the account or course level through URL, XML, or manual entry. We will use configuration via XML. First, we need to navigate into **Settings** link at the bottom of the course menu. Then navigate to the **Apps** tab which will open External Apps page. After that, on the right side of the page click on the **View App** Configurations tab and then on **Add New App**. It will open the configuration form where we need to fill Name of External Tool, Consumer Key, Shared Secret, and then choose method of configuration type (see Fig. 2). [7]

```
<?xml version="1.0" encoding="UTF-8"?>
<cartridge_basiclti_link xmlns="http://www.imsglobal.org/xsd/imslticc_v1p0"
  xmlns:blti = "http://www.imsglobal.org/xsd/imsbasiclti_v1p0"
  xmlns:ltilcm = "http://www.imsglobal.org/xsd/imsilticm_v1p0">
  <blti.title>Netpedia</blti.title>
  <blti.extensions.platform="canvas.instructure.com">
  <ltilcm.property name="privacy_level">public</ltilcm.property>
  <ltilcm.options name="course_navigation">
  <ltilcm.property name="uri">https://mitrikova.dev.cnl.sk/mywiki/index.php</ltilcm.property>
  <ltilcm.property name="enabled">true</ltilcm.property>
  </ltilcm.options>
  </blti.extensions>
</cartridge_basiclti_link>
```

Fig.2 Configuration form in Canvas needs to be filled properly by values provided by Tool Provider.

We picked Paste XML as Configuration Type. This XML should be provided by Tool Provider like Shared Secret and Consumer Key. After the configuration is successful, we have our application prepared for use in Course Navigation (see Fig. 3).

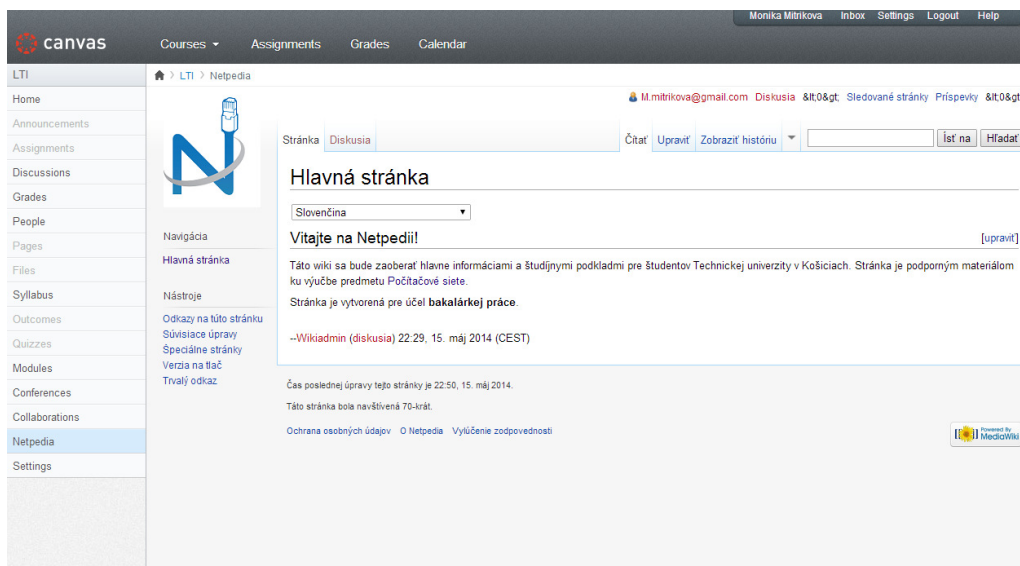


Fig.3 Netpedia as External application in system Canvas.

External application can be used for reviewing learning content and for Instructors there is also functionality that allows editing, deleting and adding content into pages.

#### IV. CONCLUSION

Implementation of the environment for creation and reviewing of learning materials in the form of LTI application for Canvas system was deployed on a school server and tested. Configuration via XML file and by entering the shared secret and consumer key has been running with no problems. This application is prepared for use in system NetAcad, which is used by Cisco Networking Academy on Technical University in Košice. Big advantage of this application is storing data on a local server, which may help for future administration. Also it will benefit instructors by simplifying content creation and adding it into their courses in a modern way. This can be very beneficial for students too, because they will have a very user friendly interface which in my opinion will ease the study of new topics.

#### ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

#### REFERENCES

- [1] Main information for developers about Canvas, available on the Internet : <<https://github.com/instructure/canvas-lms/wiki>>
- [2] Canvas REST API, available on the Internet: <<https://canvas.instructure.com/doc/api/index.html>>
- [3] Official IMS Global Learning Consortium website, available on the Internet: <<http://developers.imsglobal.org/catalog.html>>
- [4] Tutorials about LTI on the Instructure website, available on Internet: <<http://guides.instructure.com/as/2240/m/4214/1/74563-what-is-lti>>
- [5] Tutorials about LTI, available on the Internet <https://www.edu-apps.org/code.html>
- [6] Website about the Mediawiki, available on the Internet: <<http://www.mediawiki.org/wiki/MediaWiki>>
- [7] Guide by Instructure, Configuration of external applications, available on the Internet: <<http://guides.instructure.com/m/4152/l/74552-how-do-i-configure-an-external-app-for-a-course-using-xml>>

# Man-Machine rozhrania pre mobilné zariadenia v IaRS

*Karim TAFSI, Tomáš LOJKA, Iveta ZOLOTOVÁ*

Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

[karim.tafsi@student.tuke.sk](mailto:karim.tafsi@student.tuke.sk), [tomas.lojka@tuke.sk](mailto:tomas.lojka@tuke.sk), [iveta.zolotova@tuke.sk](mailto:iveta.zolotova@tuke.sk)

**Abstract**—Článok popisuje návrh a realizáciu aplikácie pre servisnú údržbu rozsiahlej vodárenskej siete. Riešenie predstavuje implementáciu mam-machine rozhrania v mobilných zariadeniach. Aplikácia je realizovaná na báze WCF service-u nasaditeľného do cloud-u, ktorý prostredníctvom použitia OPC klienta komunikuje s OPC serverom. OPC server vytvára prepojenie na PLC automat, kde je realizovaná virtuálna vodohospodárska sieť spolu s náhodným generovaním chyb vodohospodárskej siete. Generovaná chyba obsahuje informácie ako popis chyby, GPS súradnice, kde chyba nastala, popis a typ chybového zariadenia. Tieto informácie slúžia na rýchlu identifikáciu chybového zariadenia a pre jeho výmenu. Aplikácia je ošetrená aj pre výpadky spojenia a umožňuje interaktívnejšiu spoluprácu úržby s celým vodohospodárskym systémom.

**Keywords** — IaRS, mobilné zariadenia, cloud

## I. INTRODUCTION

V súčasnej dobe existuje množstvo systémov, ktoré sú rozložené na rozsiahlom území. Príkladom takéhoto systému je vodárenská sieť. Správa takýchto systémom je náročná a najmä pri potrebe realizácie rýchlej údržby. Ešte náročnejšie je realizácia údržby pri monitorovaní rozsiahlejších vidieckych vôd a odpadových zariadení, ktoré čelia rozpočtovým obmedzeniam, nemôžu si dovoliť počiatočné investície do systému SCADA/HMI. Preto cieľom práce bolo vytvoriť systém pre sledovanie porúch v rozsiahlej vodárenskej sieti. Kvôli rozsiahlosti siete je nutné, aby takáto SCADA/HMI už nebolo „privátna“, resp. realizovaná na lokálnej sieti, ale dostupná na každom mieste vodohospodárskeho systému. Tu sa nachádza miesto pre implementáciu nového technologického trendu a to cloud-u [1].

Súčasným trendom SCADA/HMI a cloudu je ich spoločná integrácia, kde výsledkom je systém zabezpečujúci dostatočný flexibilný výkon, dostupnosť a spoľahlivosť. Cloud-based SCADA/HMI, je preto veľmi atraktívnou alternatívou pre vodný priemysel a priemysel odpadových vôd. V konečnom dôsledku vytvorená aplikácia predstavuje riešenie na zníženie nákladov na kúpu, inštaláciu prevádzku a údržbu hardvéru a softvéru.

## II. ANALÝZA VYUŽITEĽNOSTI SCADA/HMI SYSTÉMOV PRE VODOHOSPODÁRSKY SYSTÉM

Rozhranie Human-machine je miesto kontaktu človeka a stroja. Úlohou man-machine rozhrania pre mobilné zariadenia je urobiť funkcionality danej technológie jasnou, prehľadnou a presne pochopiteľnou. Na to, aby bola technológia prezentovateľná pre človeka je potrebné údaje zo systému zbierať. To pokrýva funkcionality SCADA/HMI systémov [1]. V súčasnosti SCADA/HMI neumožňuje len sledovať a riadiť zariadenia a procesy, ale taktiež poskytujú širokú škálu informácií týkajúcich sa strojov a prevádzok. Tým stúpa množstvo dát, ktoré musí SCADA/HMI v rámci IaRS systému spracovávať [2]. V nami riešenom probléme bolo nutné navrhnúť riešenie pre rozsiahly vodohospodársky systém.

### A. Cloud base SCADA

Jedným z najnovších vývojov vo vzdialenom prístupe je hostovanie SCADA/HMI systému v prostredí na báze cloudu. To umožňuje užívateľom výrazne znížiť náklady a dosiahnuť väčšiu spoľahlivosť a zároveň získať možnosť vzdialene pristupovať k dátam z najrôznejších hardvérových platforiem. Okrem odstránenia nákladov a problémov spojených s hardvérovou vrstvou infraštruktúry IT, tzv. cloud-based SCADA/HMI, umožňuje používateľom zobraziť

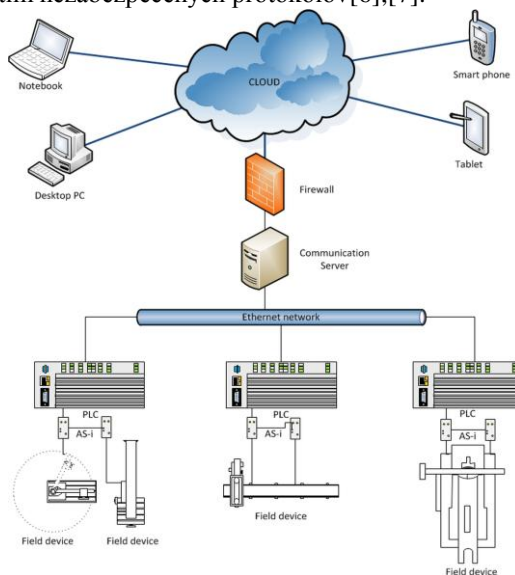
údaje na zariadení, ako sú smartfóny a tablety. Inou formou je využitie textových správ SMS či e-mailov. SCADA/HMI systém je pripojený k službe v cloudu, ktorý poskytuje vizualizáciu, podávanie správ a prístup k vzdialeným používateľom [3],[4]. SCADA/HMI je aplikovaná do cloud-u dvoma hlavnými filozofiami nasadenia:

1) Aplikácia SCADA/HMI beží priamo na mieste prevádzky, je priamo spojená s riadiacou sieťou a posielala informácie do cloudu, kde sa data ukladajú a spracúvajú. Je tu možné už realizovať analýzu nad datami (SPC) potrebnú pre správne zobrazovanie dát.

2) Celá aplikácia beží v prostredí cloudu a je vzdialene pripojená do riadiacej siete. V takejto implementácii sa nevyužíva celá funkcionálna SCADA/HMI lokálne, ale zozbierané data priamo putujú do cloud-u.

Väčšina cloud-based systémov využíva prvý spôsob. Funkcie aplikácie SCADA/HMI sú úplne izolované na riadiacej sieti, bez rozhrania cloud-u. Funkcie riadenia aplikácie SCADA/HMI sa vykonávajú len prostredníctvom riadiacej siete. Avšak aplikácia SCADA/HMI je pripojená aj k službe bežiacей v prostredí cloud, ktorá umožňuje vzdialeným používateľom nielen pristúpiť k aplikácii, ale aj využívať vizualizáciu či reportingy. Tento typ aplikácie sa často nasadzuje pri využití verejnej cloud infraštruktúry [5].

V prostredí implementácie SCADA/HMI do cloudu sa rozpráva o hrozbách zneužitia alebo aj hacknutia aplikácii bežiacich v cloud-e. V tomto smere je zvýšene nebezpečenstvo u SCADA/HMI aj použitím nezabezpečených protokolov[6],[7].



Obr. 1. Cloud base SCADA kde celá aplikácia beží na strane cloudu [8]

### III. NAVRHNUTÉ RIEŠENIE

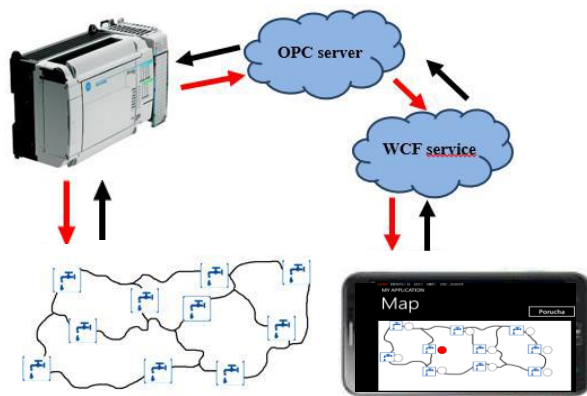
Riešenie problému si vyžiadalo komplexný analytický pohľad na aktuálnu situáciu vo oblasti spravovania vodohospodárskych zariadení. Zamerali sme sa na riešenie problémov vo vodohospodárskej oblasti, ktorá obsahuje 10 zariadení. Tieto zariadenia spravuje jeden supervízor, ktorý má informácie vo forme papierov o každom jednom zariadení a sleduje chod týchto zariadení na veline [8],[9]. V prípade poruchy jedného zo zariadení je povinný ísť k tomuto zariadeniu, a po zistení chýb, si ich spísať do papierov o zariadení. Na základe toho privolať kompetentného pracovníka, ktorý túto chybu vyrieši.

Po analýze sme vytvorili aplikáciu pre servisnú údržbu rozsiahlej vodárenskej siete prostredníctvom mobilného zariadenia. Riešenie spočíva v tom, že správca siete už nebude potrebovať množstvo papierov s informáciami, nebude musieť vysedávať na veline a taktiež nebude musieť navštevovať tieto zariadenia osobne. Štós papierov nahradí akékoľvek mobilné zariadenie, napríklad smartphone, tablet, ipad a podobne. Všetky informácie o jednotlivých zariadeniach vodohospodárskej siete sú poskytované operátorom použitím OPC servera, WCF (Windows Communication Foundation) služby-u a mobilného zariadenia. WCF služba umožňuje prepojenie mobilných zariadení a cloud-u. Tým v navrhovanom riešení získame funkcionality cloud riešenia ako je vysoká dostupnosť SCADA/HMI systému [10].

Vyžitím tohto riešenia môže operátor z akéhokoľvek miesta skontrolovať či pracujú zariadenia správne, či nastala nejaká porucha na niektorom zariadení, alebo či porucha nemala

vpliv aj na okolité zariadenia. Takato globálna diagnostika pre údržbu rozsiahlej vodárenskej siete umožňuje rýchlejši a cielenejší zásah v teréne. Tým dochádza k redukovaniu času na obnovu chodu vodohospodárskej siete.

Napríklad pri vzniku chyby na zariadení vodohospodárskej siete, je operátor túto chybu schopný na mobilnom zariadení vidieť, nájsť si presne GSP súradnice umiestnenia zariadenia, dôvod chyby, vážnosť poruchy. Pri dokončení opráv môže si overiť stav zariadenia v systéme a následne potvrdiť zistenú chybu. Tým prehlási, že na zariadení už nebude vykonávaný žiaden zásah a zariadenie je schopné prevádzky.



Obr. 2. Zjednodušené znázornenie funkcionality navrhnutého riešenia

#### IV. REALIZOVANÉ RIEŠENIE

Aplikácia je vytvorená v prostredí Microsoft Visual Studio 2010 v jazyku C# a RSLogix 5000. Vo Visual Studio 2010 je vytvorená aplikácia pre mobilné zariadenia, ktoré pracujú pod operačným systémom Windows Phone. Na simuláciu mobilného zariadenia bol použitý Windows Phone Emulator. Simulácia situácie rozsiahlej vodárenskej siete je znázornená na Obr.2.

##### A. Popis implementovaného riešenia

Na začiatku aplikácie je používateľ mobilného zariadenia vyzvaný, aby zadal prihlasovacie meno a heslo pre prihlásenie sa do systému. Pre testovanie a beh rozsiahlej vodárenskej siete v laboratóriu bol využitý PLC automat Compact Logix L23E. Na ňom bol vytvorený program, ktorý simuluje chod a náhodne generovanie chýb vodohospodárskej siete. Na komunikáciu s PLC bol použitý OPC server. Vznik chyby detekuje OPC server a vytvorí udalosť na WCF službe. S WCF službou interaguje mobilné zariadenie. WCF služba zabezpečuje pre naše zariadenie signál, spája s OPC serverom, podobne ako operátor pre naše mobilné telefóny. Zariadenie, na ktorom bola detekovaná chyba sa na mape zobrazí červenou farbou (Obr. 3). Údaje o tomto zariadení sa zobrazia následne po kliknutí na toto zariadenie ako aj detailný popis chyby. Tieto zariadenia majú svoje špecifické údaje ako sú: ID zariadenia, názov riadenia, lokalita, v ktorej je umiestnené, názov zariadenia, stav v akom sa nachádza (online/offline), prioritu vyšetrenia, GPS súradnice, informáciu, či bol na ňom error potvrdený, alebo nie a detailný popis chyby (získany z diagnostiky zariadenia).

Pri výbere zariadenia sa načítajú údaje o zariadení do „faceplate“ obrazovky, ktorá slúži ako šablona pre každé zariadenie vo vodohospodárskej sieti.

##### B. Ošetrovanie výpadkov komunikácie OPC, WCF a mobilného zariadenia

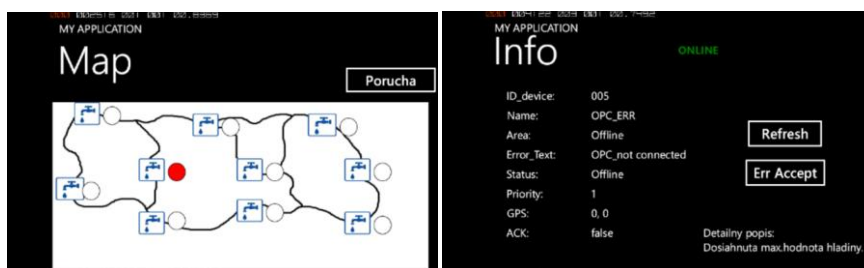
Počas behu aplikácie môže dochádzať k výpadkom aplikácie, či dlhodobejšej strate spojenia na strane zariadenia. V prípade, že dôjde k strate spojenia WCF služby z cloudu s OPC serverom, tak WCF služba poskytne údaje offline režime. Inak povedané poskytne posledné dostupné informácie, pričom zašle správu zariadeniam, že stratil spojenie so zariadeniami v technologickej sieti.

V prípade výpadku spojenia mobilného zariadenia s WCF službou, dôjde zobrazeniu stavu OFFLINE v pravej časti obrazovky. V oboch prípadoch, operátor mobilnej aplikácie môže pracovať s datami, ktoré sú aktuálne dostupné ako napríklad s detailným popisom chyby na všetkých chybových zariadeniach. Operátor môže po ukončení opráv vykonať súhlas



s ukočením prác (potvrdiť poruchu). Jeho správanie, ktoré vykonal sa uchováva v pamäti mobilného zariadenia. Pri obnovení spojenia ukony ktoré operátor vykonal sa automaticky odošlú WCF služba-u.

Na Obr.3 sú znázornené dve obrazovky



Obr. 3. Znáozornenie obrazoviek aplikácii

## V. ZÁVER

Navrhnutá a realizovaná aplikácia je príkladom, ako môžu mobilné zariadenia napr. pomôcť pri údržbe rozsiahlej vodárenskej siete. Presunom SCADA/HMI do prostredia cloud riešení môžu poskytovatelia a používatelia systémov SCADA/HMI výrazným spôsobom znížiť náklady, dosiahnu vyššiu spoľahlivosť a rozšíriť funkcionality celého systému. Okrem eliminácie nákladov a problémov spojených s hardvérovou a IT infraštruktúrou SCADA/HMI systému na báze cloud, umožňujú používateľom zobrazovať údaje na zariadeniach, ako sú inteligentné mobilné telefóny či tablety, a využívať aj SMS či e-mail.

## ACKNOWLEDGMENT

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: **Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií**, kód ITMS: **26220220182**, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja (60%). Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ. Príspevok bol tiež podporený grantom z KEGA - 021TUKE-4/2012 (40%).

## REFERENCES

- [1] A. Daneels, S. Wayne, "What is SCADA." In: *International Conference on Accelerator and Large Experimental Physics Control Systems*. 1999.
- [2] M. Karacor, E. Ozdemir, „Mobile phone based SCADA for industrial automation“, in: *ISA Trans*, 45 (2006) Volume 45, Number 1, January 2006, pages 67–75
- [3] C. Alcaraz, I. Agudo, D. Nunez, J. Lopez, “Managing Incidents in Smart Grids à la Cloud”, *Cloud Computing Technology and Science (CloudCom)*, 2011 IEEE Third International Conference, pp.527-531, Nov. 29 2011-Dec. 1 2011.
- [4] L. Youseff, M. Butrico, D. Da Silva, "Toward a Unified Ontology of Cloud Computing", *Grid Computing Environments Workshop, 2008, GCE '08*, pp.1-10, 12-16 Nov. 2008, doi: 10.1109/GCE.2008.4738443.
- [5] G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, W. Vogels, “Dynamo: amazon's highly available key-value store”, in *SOSP '07*. New York, NY, USA: ACM, 2007, pp. 205-220.
- [6] C. Alcaraz, G. Fernandez, R. Roman, A. Balastegui, J. Lopez, “Secure Management of SCADA Networks”, *New Trends in Network Management, CEPIS by Novática (ATI, Spain)*, vol. 9, no. 6, pp. 22–28, December 2008.
- [7] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, E. G. Im, Z.Q Yao, B. Pranggono, H.F. Wang, "Man-in-the-middle attack test-bed investigating cyber-security vulnerabilities in Smart Grid SCADA systems," *Sustainable Power Generation and Supply (SUPERGEN 2012)*, pp.1,8, 8-9 Sept. 2012.
- [8] T. Lojka, I. Zolotová, „Improvement of Human-plant Interactivity via Industrial Cloud-based Supervisory Control and Data Acquisition System“, in: *IFIP Advances in Informatics and Communication Technology*, submitted for publication.
- [9] J. Liguš, I. Zolotová, P. Karch, J. Ligušová, “Information and control system of traverse and its integration into cybernetic centre”, *Electronics and electrical engineering*, no. 6 (2010), p. 147-152, June 2010, Print ISSN: 1392-1215, Online ISSN: 2029-5731.
- [10] J. Lowy, *Programming WCF Services: Mastering WCF and the Azure AppFabric Service Bus*. O'Reilly Media, Inc., 2010.

# Kontrolný video dohľadový systém pre monitorovanie cestnej premávky v tuneli Branisko

<sup>1</sup>Lukáš HILLAY, <sup>1</sup>Luboš OVSEŇÍK, <sup>1</sup>Ján TURÁN

<sup>1</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>lubos.ovsenik@tuke.sk, <sup>1</sup>jan.turan@tuke.sk

**Abstrakt** — Predložená práca na základe dostupných prostriedkov a doterajších skúseností približuje fungovanie video dohľadového systému (VDS), ktorý sa používa v tuneli Branisko. Zároveň poukazuje na jeho nastavenie, kalibráciu i jeho samotné fungovanie. Našou ambíciou je porovnať zmeny, ktoré nastali v tuneli a ich dopady na fungovanie VDS.

**Kľúčové slová** — detekčná zóna, monitorovanie cestnej premávky, systém Autoscope, video dohľadový systém

## I. ÚVOD

Monitorovanie pomocou kamerového systému pokladáme dnes za úplnú samozrejmosť. Inteligentné monitorovacie systémy tvoria základné prvky bezpečnosti. Zabezpečenie po hardvérovej stránke je rovnako dôležité, ako po stránke softvérovej. Vždy závisí od oblasti implementácie a požiadaviek na systém. Video monitoring dnes patrí k základným prvkom bezpečnosti. S kamerovými systémami sme sa stretávali dávnejšie iba v bankách či nákupných centrách. Vývoj napreduje míľovými krokmi. Dnes nachádzajú široké uplatnenie v mnohých oblastiach. Vývojári navrhujú inteligentné video dohľadové systémy, ktoré dokážu pracovať plne automaticky bez riadenia človekom. Ich výhodou je bezpochyby včasné varovanie a tým aj prevencia pred nebezpečnými situáciami.

Inteligentné vizuálne zabezpečovacie systémy sa v reálnom čase zaoberajú monitoringom perzistentných a prechodných objektov v rámci špecifického prostredia. V cestnej premávke dokážu takéto systémy upozorniť vodičov na možné riziká, medzi ktoré možno zaradiť diaľnicu, tunel, či horské prechody [1-6].

Zvyšujúci sa dopyt po bezpečnosti vedie k rastúcej potrebe VDS v mnohých oblastiach. V poslednej dobe sa stále častejšie hovorí o monitoringu z hľadiska bezpečnosti. Jedným zo základných delení video dohľadových systémov je na vonkajšie a vnútorné.

- *Vnútorné*: kamerový systém sníma interiér. Snímané pozadie je do istej miery statického charakteru. Snímaná scéna sa mení iba minimálne, menia sa iba objekty v scéne.
- *Vonkajšie*: sú analogicky použité v exteriéry. Tieto systémy vyžadujú vyšší stupeň flexibility. Vplyvy vonkajších činiteľov i samotných objektov majú charakter dynamického pozadia (detekcia vozidiel, chodcov, požiaru, atď.). Snímaná scéna je veľmi rozmanitá je potrebná zložitejšia detekcia, ako pri vnútorných systémoch.

Ďalším kritériom je delenie na základe snímaných objektov. Z hľadiska video dohľadových systémov rozlišujeme:

- *Snímanie statických objektov*: zabezpečenie budovy, zabezpečenie súkromných pozemkov, zabezpečenie objektov, ktoré sa vzhľadom na snímanú scénu nehýbu.
- *Snímanie dynamických objektov*: ich základnou vlastnosťou je pohyb, a teda snímaná scéna sa neustále mení. Pohyblivé objekty nájdeme všade od motorových vozidiel až po lode či vlaky. Video dohľadové systémy z hľadiska použitia v doprave môžeme použiť v oblastiach:
  - *Letecká doprava*: patrí sem najmä bezpečnosť na letisku, detekcia batožiny, detegovanie hľadaných a nebezpečných osôb.

- *Námorná doprava*: snímanie prístavov, snímanie lodí, detekcia kolízií.
- *Železničná doprava*: snímanie železničných pricestí, snímanie nebezpečných úsekov, železničné výhybky a semaforey.

- *Cestná doprava*: snímanie križovatiek, snímanie diaľnic a tunelov.

Z hľadiska použitia v mestskej infraštruktúre, kde sú husto zaľudnené a zastavané oblasti je potrebný monitorovací systém v týchto dvoch hlavných oblastiach:

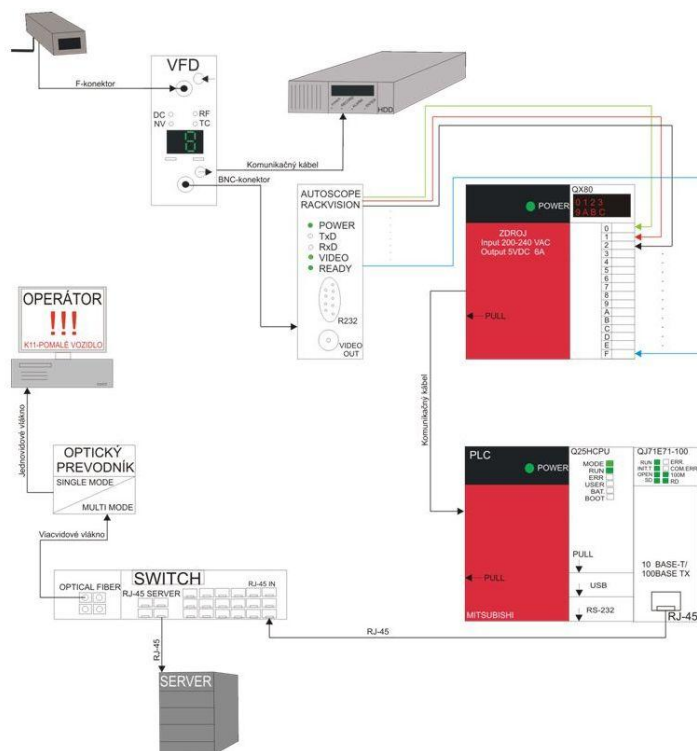
- *Verejné miesta*: patria sem banky, supermarkety, domy, parkoviská, rušné ulice, oblasti s vysokou kriminalitou, vandalizmus - grafity.
- *Verejné udalosti*: rôzne športové, kultúrne či inak spoločensky zamerané udalosti, kde sa pohybuje množstvo cudzích a potenciálne nebezpečných ľudí. V poslednej dobe sa čoraz častejšie hovorí o hrozbe terorizmu [2-6].

## II. VIDEO DOHLADOVÝ SYSTÉM V TUNELI BRANISKO

V cestnej premávke predstavuje tunel vysoké riziko vzniku nebezpečných situácií. V tuneli Branisko sa využíva navyše obojsmerná prevádzka, čo zvyšuje nebezpečenstvo dvojnásobne. Dĺžka tunela je približne 5 kilometrov, t.j. pomerne dlhý úsek uzavretej vozovky. Je preto úplnou nevyhnutnosťou mať komplexný monitorovací systém. Od merania ovzdušia, cez kamerový systém až po osvetlenie tunela. Pracuje tu mnoho systémov vedľa seba. Naša práca sa zaoberá video dohľadovými systémami a preto sa zameriame na systém Autoscope v tuneli Branisko [1].

### A. Systém Autoscope

Na Obr. 1 sa nachádzajú jednotlivé segmenty VDS Autoscope. Začína vstupný článkom, ktorý predstavuje kamera na snímanie cestnej premávky. Nasleduje video - karta, ktorá dokáže spracovať signál z kamery. Okrem toho umožňuje zálohu všetkých nahrávok na externé HDD. Z video - karty pokračuje signál do Autoscopovej karty, ktorá je hlavná časť hardvérovej výbavy. Jej výstupy sú napojené na binárne karty, z ktorých sa signál ďalej spracováva cez PLC procesor. Jeho výstupom je rozhranie RJ-45, ktoré umožňuje pripojenie na switch, odtiaľ je možné pripojenie na server. Nastáva zmena prenosového média z metalického na optické. Následne sa používa optický prevodník, ktorý umožňuje zmenu optického vlákna z multividového na jednovidové. Signál pokračuje po optickom vedení až na riadiace stredisko.

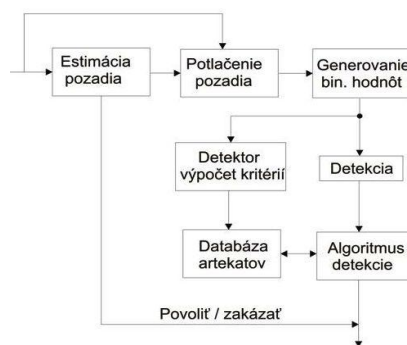


Obr. 1 Hardvérová schéma zapojenia systému Autoscope v tuneli Branisko.



### B. Algoritmus Autoscope

Algoritmus Autoscope je znázornený na Obr. 2. Prvý stupeň spracovania je zodpovedný za zníženie objemu spracovaných údajov. Obsahuje blok „Estimácia pozadia“, ktorého úlohou je odhad pozadia detektora. Nato využíva FIR filter s konečnou impulzovou odozvou. Nasleduje blok „Potlačenie pozadia“, ako už zo samotného názvu vyplýva, jeho úlohou je potlačenie pozadia na úkor presnejšej detekcie vozidla. Ďalším stupňom spracovania je extrahovanie funkcií a generovanie skupiny binárnych hodnôt. Detektor detekuje prítomnosť vozidla na základe určitých kritérií, ktorých výpočet uskutočňuje blok „Detektora“. Tento blok môže mať k dispozícii, ale nemusí (závisí od verzie Autoscope) aj databázu artefaktov respektíve rušivých prvkov, ktoré uľahčujú samotný výpočet kritérií. Posledným blokom je „Algoritmus na detekciu vozidiel“. Jeho výstupom je rozhodnutie či tam je, alebo nie je prítomné vozidlo. Nakoniec je potrebné spomenúť všeobecný prístup k eliminácii niektorých z hlavných artefaktov ako sú tieň, odrazy vozidiel na chodníku, atď. Pre pevné objekty ako sú tieň, sú tieto kompenzované v závislosti od potlačenia pozadia. U pohybujúcich sa tieňov je eliminácia založená na upravení zisku, alebo skreslení javov [1].

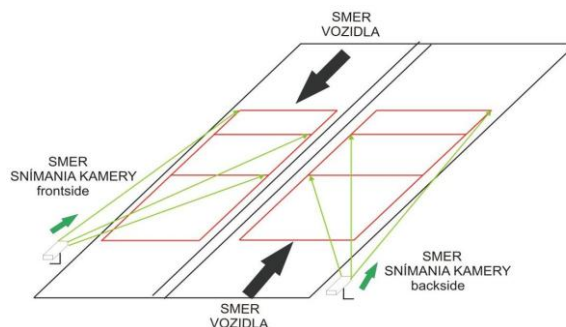


Obr. 2 Algoritmus systému Autoscope.

### C. Základné delenie kamier pre VDS Autoscope

V tunelovej „rúre“ existuje niekoľko typov kamier. Ich umiestnenie a používanie je dané trajektóriou tunelovej rúry. Kvôli bezpečnosti existuje v tuneli niekoľko núdzových odstavných plôch hovorovo nazývané zálivy. Slúžia na odstavenie vozidla v prípade poruchy, nečakanej situácie, požiaru, atď. Nachádza sa v nich aj núdzový telefón inak nazývaný SOS - ka. Primárne rozdelenie je na základe snímaných častí a to na kameru:

- Zálivová: sníma záliv – núdzovú odstavnú plochu. V tunelovej rúre je odstavná plocha na jednej strane vozovky, alebo na oboch stranách vozovky – obojstranný záliv.
  - Nezálivová frontside: sníma prichádzajúce vozidlá v priamom smere prichádzajúce čelom ku kamere.
  - Nezálivová backside: sníma vozidlá prichádzajúce v priamom smere odzadu ku kamere.
- Pre názornejšie pochopenie nám posluží Obr. 3, kde vozidlá prichádzajú v smere šípkov umiestnených na vozovke.



Obr. 3 Základné delenie kamier v tuneli Branisko pre VDS Autoscope.

### III. EXPERIMENTÁLNE OVERENIE SYSTÉMU AUTOSCOPE

Nastavovanie a testovanie systému Autoscope sa uskutočňuje vždy pri odstávke tunela. Tunel je uzavretý každý rok v máji a v októbri. V tomto období je vymedzený čas na údržbu a testovanie. Kamery sa čistia počas údržby. Robí sa servis na všetkých zariadeniach systému. Odstraňujú sa vzniknuté problémy: chybná kamera, karta a podobne. Po servise je potrebná kontrola nastavení a testovanie správnej funkčnosti systému. Je nutné poznamenať, že jednotlivé skúšky sa vykonávajú na každej kamere umiestnenej v tunelovej rúre. Celkový čas detekcie musí spĺňať požadovaný limit. Naša práca zahŕňa výsledky meraní od roku 2011 do roku 2013. Počas týchto troch rokov došlo v tuneli k nasledujúcim zmenám, ktoré mali a majú vplyv na funkčnosť systému Autoscope:

- Rok 2011: tunel v pôvodnom stave.
- Rok 2011-2012: osadenie LED diódového osvetlenia. Na okraji vozovky boli nainštalované biele a oranžové LED osvetlenie, ako aj zelené LED pri únikových východoch. Ich osadenie bolo za účelom zvýšenia bezpečnosti v tuneli.
- Rok 2013: prebiehalo maľovanie bočných strán tunela. Náter by mal byť ľahko umývateľný a nenáročný na údržbu, no spôsobuje odrazy svetiel či už prechádzajúcich automobilov alebo osvetlenia tunela.

#### A. Základné skúšky

- Udalosť č. 1 – Zastavené vozidlo

Táto skúška spočíva v koordinácii operátora a vozidla v tuneli. Pri správnej činnosti je udalosť vygenerovaná najneskôr do 45 s. Reálne je tento čas ovplyvňovaný mnohými faktormi: veľkosť vozidla; vzdialenosť medzi vozidlom a predmetnou kamerou; úroveň osvetlenia v tunelovej rúre; povrch vozovky hlavne vplyv vody; majáky – vozidlá s právom prednostnej jazdy, údržba; farba vozidla, chrómované časti – odraz, lom, disperzia svetla; atď. Po vygenerovaní príslušnej udalosti sa vozidlo presunie do zorného uhla nasledujúcej kamery, kde sa zopakuje celá procedúra. Ak sa daná udalosť nedetegovala pri niektorej kamere, musí vozidlo opustiť jej zorné pole a zopakovať procedúru znova.

Výsledky meraní pre modelovú situáciu zastavené vozidlo pre rok 2011 sú nasledujúce (Tab. 1). Celkový priemerný čas všetkých kamier (K03 - K71) dosahoval hodnotu  $t_1 = 36,44643$  s. Ide o priemerný údaj, to znamená, že niektoré kamery reagovali skôr iné zasa neskôr. V tomto smere ide o určité skreslenie.

Pre rok 2012 bol celkový priemerný čas rovný  $t_2 = 33,67857$  s. Čo reprezentuje celkový pokles doby detekcie o 8 %. Inak povedané systém Autoscope po nastaveniach a kalibrácií dokázal detegovať zastavené vozidlo o 8 % skôr v porovnaní s predchádzajúcim rokom. V tomto období boli osadené LED osvetlenia v tuneli Branisko.

Pre rok 2013 bol celkový priemerný čas rovný  $t_3 = 35,44643$  s. To reprezentuje zlepšenie na úrovni 3 % v porovnaní s rokom 2011. V tomto roku bol nanosený bočný náter na steny v tunelovej rúre.

Tabuľka 1  
Celkové výsledky meraní z rokov 2011, 2012, 2013 pre „Udalosť č. 1 – zastavené vozidlo“

Rok	2011	2012	2013
Priemerný čas všetkých kamier [s]	36,44643	33,67857	35,44643
Porovnanie v [%]	100 %	-8 %	-3 %
Podmienky tunel Branisko	Bez LED	LED	LED + Náter
Označenie	$t_1$	$t_2$	$t_3$

- Udalosť č. 2 – Pomalé vozidlo

Rýchlosť vozidla je pri tejto skúške 25 – 30 km/h. V smere Levoča - Prešov, bude pomalá rýchlosť vozidla detegovaná len kamerami nachádzajúcimi sa v tom jazdnom pruhu, okolo ktorých sa vozidlo pohybuje (teda nie kamerami na opačnej strane vozovky). V smere Prešov – Levoča, bude pomalá rýchlosť vozidla detegovaná len kamerami nachádzajúcimi sa v tom jazdnom pruhu, okolo ktorých sa vozidlo pohybuje (nie kamerami na opačnej strane vozovky).

Výsledky meraní pre modelovú situáciu pomalé vozidlo pre rok 2011 sú nasledujúce. Celkový priemerný čas všetkých kamier (K03 - K71) dosahoval hodnotu  $t_1 = 32,12766$  s. Pre rok 2012 bol celkový priemerný čas rovný  $t_2 = 24,56574$  s. Čo reprezentuje celkový pokles doby detekcie o 23 %. Pre rok 2013 bol celkový priemerný čas rovný  $t_3 = 27,76596$  s. Čo

reprezentuje celkový pokles doby detekcie o 13 % (Tab. 2).

Tabuľka 2  
Celkové výsledky meraní z rokov 2011, 2012, 2013 pre „Udalosť č. 2 – pomalé vozidlo“

Rok	2011	2012	2013
Priemerný čas všetkých kamier [s]	32,12766	24,56574	27,76596
Porovnanie v [%]	100 %	-23 %	-13 %
Podmienky tunel Branisko	Bez LED	LED	LED + Náter
Označenie	t <sub>1</sub>	t <sub>2</sub>	t <sub>3</sub>

- Udalosť č. 3 – Vozidlo v protismere

Skúška spočíva v prejení tunelovej rúry z východného portálu na západný portál teda zo smeru Prešov –Levoča a naopak. Motorové vozidlo prechádza vozovku v tuneli v protismere. Rýchlosť vozidla je minimálne 50 km/h a viac. V tuneli je dovolená maximálna rýchlosť 80 km/h.

- Udalosť č. 4 - Chyba kamery a vychýlenie kamery

Túto skúšku vykonávame pomocou prerušenia kábla medzi kamerou a Autoscopom. Nasimulovanie prípadného rozbitia kamery, alebo stratou napájania kamery bude realizované pomocou vypnutia PAL video signálu z video vstupu príslušnej karty. Signál bude prerušený prostredníctvom BNC konektorovej spojky umiestnenej v rozvážači Autoscope.

- Udalosť č. 5 – Status karty

Status karty môže nadobúdať 2 stavy a to „ON“ a „OFF“. O zapnutom stave hovoríme ak je software pre detekciu bez poruchy. Opačný stav nastane, keď je porucha pri spracovaní software (porucha na elektronike karty) výstup bude deaktivovaný, t. j. bude v stave „OFF“. Status karty je teda udalosť, ktorá bude generovaná len pri poruche karty, čo nie je možné odskúšať bez fyzického poškodenia karty. Status karty teda možnosť overenia či je výstup aktívny alebo nie, je možný len pri reštarte karty. Vtedy sa deaktivujú všetky výstupy karty aj status karty.

- Udalosť č. 6 – Intenzita dopravy

Všetky kamery umiestnené v tunelovej rúre smerujúce na vozovku detegujú udalosti č. 1, 2, 3. Existujú však aj kamery, ktoré sa nachádzajú pred vjazdom do tunela, ako z jednej tak i druhej strany. Ich účel nie je detegovať stojaté, pomalé či protiídúce vozidlo, ale poskytujú informácie o dopravnej záťaži v tuneli. Výsledný počet vozidiel za hodinu sa prepočítava koeficientom na opacitu.

Opacita je zariadenie v tuneli, ktoré umožňuje merať množstvo exhalátov v tunelovej rúre. Na základe jej meraní sa reguluje výkon vzduchotechniky. Pri automatickom riadení vzduchotechniky je vetranie riadené vždy podľa najnepriaznivejšej hodnoty vybranej z opacity, CO<sub>2</sub> a dopravnej záťaže. Presnosť určenia dopravnej záťaže Autoscopom je v intervale  $\pm 5$  %. Táto presnosť postačuje na riadenie vzduchotechniky, ale nepostačuje na presnú identifikáciu počtu vozidiel v tunelovej rúre v prípade požiaru.

### B. Zhodnotenie základných skúšok

Z výsledkov za uplynulé tri roky možno vyvodit' záver, že nastalo celkové zníženie odozvy systému na „Zastavené“ a „Pomalé vozidlo“. Tento záver však nie je celkom jednoznačný. Výsledky pochádzajú z experimentálnych meraní. Merania sa uskutočňujú mimo riadnej prevádzky. Tunel je odstavený a hustota vozidiel je veľmi nízka v porovnaní s bežným chodom. Po zavedení LED osvetlenia a bočného náteru sa podstatne zvýšil počet falošných alarmov. Rápidne zvýšenie bolo hlavne počas nočnej prevádzky, kedy intenzita osvetlenia klesla. Tým nechceme povedať, že merania sú úplne zbytočné. Na zariadeniach sa robí pravidelný servis a manipuluje sa s nimi. Po stránke správneho fungovania je to nevyhnutné.

Osadenie LED osvetlenia prinieslo na jednej strane o čosi svetlejšiu scénu, ktorú snímajú kamery. To mohlo viesť aj k spomínanému zníženiu odozvy systému o 8 % pri zastavenom a o 23 % pri pomalom vozidle. Na strane druhej osadenie osvetlenia prinieslo množstvo falošných alarmov, ktoré znižujú efektívnu funkčnosť celkového systému. Po aplikovaní náteru sa podmienky ešte zhoršili. Celková odozva systému klesla na úroveň 3 % pri zastavenom vozidle a 13 % pri stojacom vozidle v porovnaní s rokom 2011.

Treba si uvedomiť, že maximálna dovolená rýchlosť v tuneli je 80 km/h. Čo reprezentuje 22,2 m/s. Vozidlo prejde za jednu sekundu 22 metrov. Celková odozva systému je v ideálnom prípade rovná 45 sekundám. Za spomínaní interval prejde vozidlo približne 990 metrov. Komplexné skúšky boli teda „Zastavené vozidlo“, „Pomalé vozidlo“ a „Jazda v protismere“.

Zastavené vozidlo, alebo akékoľvek stojace vozidlo v tuneli je veľmi nebezpečné. Preto operátori na riadiacom stredisku pokladajú práve tento typ detekcie za najdôležitejšiu. Z používateľského hľadiska by kládli vyšší dôraz na zníženie doby odozvy zastaveného vozidla, ako napríklad jazda v protismere. Z hľadiska funkčnosti by to bolo do istej miery možné. Zväčšením detekčných zón by sa celkový čas detekcie zastaveného vozidla o čosi znížil, ale detekcia jazdy v protismere by ale nefungovala, alebo fungovala s obrovským oneskorením. Takže terajšie nastavenie je určitým kompromisom.

Pomalé vozidlo je detegované pri rýchlosti okolo 30 km/h. Táto rýchlosť je skutočne veľmi pomalá v porovnaní s maximálnou dovolenou. V prípade poruchy odstaví vodič svoje vozidlo v núdzovom zálive. V praxi systém Autoscope skôr deteguje zastavené vozidlo, ako pomalé. Je to spôsobené hlavne tým, že pri stojacom vozidle nastáva okamžitá skoková zmena v snímanej detekčnej zóne. Pri pomalom sa cez detekčnú zónu vozidlo iba mihne. Je lepšie ak stojace vozidlo systém deteguje, ako pomalé a až následne stojace, ako by nedetegoval nič.

Jazda v protismere: ak vozidlo prejde určitou časťou do protismeru systém Autoscope to vyhodnotí v rozpätí 30 – 40 s. Za tento časový interval vozidlo prejde približne 20 % celkovej dĺžky tunela teda asi 1 km. Hlásenia sú skôr informatívneho charakteru a upozorňujú operátora na danú skutočnosť s podstatným oneskorením.

### C. Špecifické príklady skúšok

Tieto situácie v tuneli môžu nastať a fungovanie systému Autoscope ovplyvňuje napr.:

- Núdzové odstavenie vozidla v zálive: faktor, ktorý vplýva na reakčný čas je umiestnenie vozidla v detekčnej zóne – jeho pozícia.
- Zmena podmienok v tuneli (LED, náter).
- Vplyv počasia – dážď, slnko.
- Vozidlá s majákom.
- Denný a nočný režim osvetlenia v tuneli.

## IV. ZÁVER

Každá technológia dokáže pracovať najefektívnejšie vtedy, ak je implementovaná v prostredí pre ktoré bola navrhnutá. Systém Autoscope je veľmi perspektívna technológia. Tak prečo nefunguje tak bezproblémovo aj v tuneli Branisko? Dôvod je jednoznačný. Pri návrhu tohto systému sa počítalo so stavbou dvoch tunelových rúr. Pričom v každej z nich by bola iba jednosmerná prevádzka. Kamery mali byť umiestnené na strope a snímať vozovku „iba“ v jednom smere. Systém Autoscope by teda dokázal pracovať oveľa efektívnejšie, ak by bol pracoval za takých podmienok, pre aké bol navrhnutý. Terajší stav je teda iba určitým kompromisom a pri obojsmernej prevádzke nie je možné dosiahnuť lepšie výsledky.

## POĎAKOVANIE

Táto práca bola podporovaná Kultúrnou a edukačnou grantovou agentúrou (KEGA) Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky na základe projektu č. „063TUKE-4/2013 - Použitie diaľkovo ovládaného optického vláknového refraktometra vo výučbe“ a projektu č. „006TUKE-4/2014 - Experimentálny model PON siete TUKE a jeho použitie vo výučbe“. Táto práca je tiež výsledkom zmluvy „APVV-0025-12 - Predchádzanie vplyvu stochastických mechanizmov vo vysokorýchlostných plne optických sieťach“ podporovanej Agentúrou na podporu výskumu a vývoja (APVV) Slovenskej republiky.

## LITERATÚRA

- [1] G. Panos Michalopoulos, „Vehicle detection video through image processing: the Autoscope system“, *Vehicular Technology, IEEE Transactions on*, Feb 1991, vol. 40, no. 1, pp. 21-29.
- [2] Video Surveillance Systems (CCTV), <[http://vds.de/fileadmin/vds\\_publicationen/vds\\_5473en\\_web.pdf](http://vds.de/fileadmin/vds_publicationen/vds_5473en_web.pdf)>.
- [3] Vehicle Detection under Day and Night Illumination, <<http://imabelab.ing.unimore.it/imabelab/publicazioni/ia1999.pdf>>
- [4] Vehicle Detect And Monitor Techniques For Intelligent, <[http://www.iosrjournals.org/iosr-jmce/papers/sicete\(civil\)-volume4/48.pdf](http://www.iosrjournals.org/iosr-jmce/papers/sicete(civil)-volume4/48.pdf)>.
- [5] Intelligent Video Analytic Surveillance System, <<http://www.ass-asecurity.com>>.
- [6] Programmable Controllers MELSEC <<http://www.mitsubishielectric.com/fa/products/cnt/plc/>>.

# Automated System for Operations Management and Crisis Events for Smart Home Application

<sup>1</sup> Michal TIRPÁK, <sup>2</sup> Michal KOVALČÍK, <sup>3</sup> Peter FECILAK, <sup>4</sup> František JAKAB

<sup>1,2,3,4</sup> Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

<sup>1</sup> michal.tirpak@student.tuke.sk, <sup>2</sup> michal.kovalcik@tuke.sk, <sup>3</sup> peter.fecilak@tuke.sk,  
<sup>4</sup> frantisek.jakab@tuke.sk

**Abstract** — This paper deals with intelligent houses and solutions that enable automatization of family house or block of flats. The paper analyses some of commercially available solutions and presents their advantages and disadvantages. The analytical part describes possibilities of smart grids and their importance connected with the energy delivery. The main goal is a proposal and implementation of system which allows to control operation and critical incidents in an intelligent house. The system also enables observation of devices and sensors conditions using the application developed for that purposes. Lights, heating and other parts are controlled according to the information received from sensors and managed by a programmable logic unit.

**Keywords** — Intelligent house, Smart grid, Smart meter

## I. INTRODUCTION

In the last decades technological progress has marked a big leap. We meet with modern technologies almost everywhere. From boarding computers in cars, smart phones and tablets to complicated systems that manage factories, towns or whole countries. This progress affects also houses. It is more impressive these days because hundred years ago there was not an electricity in the most of Slovakia and today we have got smartphone or tablet and we are able to listen to our favorite song or control cursor on a screen. Modern technologies enable us to do almost unbelievable things. They move safety and comfort of living to higher level. We are able to monitor our consumption of energy in the house, control lights, windows or heating. By means of multimedia centers also ordinary house may turn to private cinema or disco. Using camera security system we are able to watch who is located where in the house [1].

## II. COMMERCIALY AVAILABLE SOLUTIONS OF INTELLIGENT HOUSES

There are many commercially available solutions. In Slovakia and in Czech Republic is the public awareness about topic of intelligent houses very low. People know very little about brands and producers. They do not have a general knowledge about prices and they do not trust in automatization of house. This situation is also caused by distributors themselves, because they offer intelligent house instead of selling solutions or concrete brands. Intelligent house may be available for 3000 euros and however also for 30 000 euros – but where is the difference? Probably it is given by fact that every producer offers something else under the term of intelligent house. For some of them it means basic control of lights, blinds, heaters and smart use of television set and all of that thanks to classic systems build on computers with operating system Windows [2].

For others it means complex system of automatization built after many years of constant improvement. These solutions use reliable and robust factory of computers and except for already mentioned functions they can control camera security system, measurement and regulation of the whole object. They can guarantee high-quality sound, image or voice distribution and distant management. From that reason there exist wide range of prices and quality [2] [3].

### III. SMART METER

Under the term of "Smart meter" we usually mean device determined to record the consumption of electric energy, such device is sending information in specific intervals to administrator for monitoring purposes. On the basis of these information we are able to establish when and how much energy is spent and distributive companies can regulate prices politics according to that. Regular consumer is able to monitor not only recent consumption on smart meter but also how much energy he spends and in which part of the day is consumption the highest.

In the Figure 1 is depicted bidirectional communication between device and the central monitoring system of energy distributor realized via internet connection. Advanced measuring utility using smart meters differs from ordinary automatic measuring because communication with device is bidirectional [14].



Fig. 1 Smart meter

Smart meters are usually mentioned in connection with measuring of consumption of electric energy but this device might also be a meter of water or gas consumption. Compared to older devices, smart meters offer data in the real time or almost real time together with data about quality of energy and its potential failures. The benefit for users might be a possibility to sell electric energy back to the distribution network or an observation of history of recordings [15].

Providers of electric energy have been searching for resources and possibilities of comparing the consumption of electric energy with its production and how to purposefully fill the demand. Ordinary measuring devices have been able to measure only total consumption and has not been able to monitor when the energy has been used. Smart meters offer these data specified for particular place of consumption of electric energy and this enables distributors to determine different prices for consumers depends on period of a day and season. Smart meters also enable monitoring of voltage in the network and potential harmonic distortion which speeds up the diagnostics of problems with delivered energy [15] [16] [17].

Companies producing energy claim that from consumers' perspective smart metering offers advantages for houses, e.g.:

- a) elimination of estimated expenses that are source of complains of consumers,
- b) smart meters as a tool for help to consumers to improve management of energy consumption – they are able to get actual information about their consumption on display.

Producers of energy hope that setting of different prices according to daily period of consumption (highest in the time of highest consumption) they will enable change habits of customers and temporarily postpone the need of making additional sources (power stations) or at least to reduce the need of taking energy away from financially expensive sources.

It is important to mention that smart meters might be a part of smart grid but they are not able to create it by themselves [18].

### IV. SMART GRID

Smart grid shown in Figure 2 is electric network that uses information-communication technologies to acquire information about behavior of customers and suppliers. The aim of smart grid is to increase automatically effectiveness, return ability, economy and sustainability of production and distribution of electric energy as well as other energies. Introduction to that technology also involves fundamental developing of energetic services. Intelligent networks guarantee that renewable sources of energy can be integrated easily to network because they enable bidirectional energy flow – from producers to consumers and vice versa (orange line) and bidirectional flow of communication (blue line).

Whereas in conventional network supply only consumption is covered, intelligent networks are able to manage consumption according to availability of electric energy in the network [19].

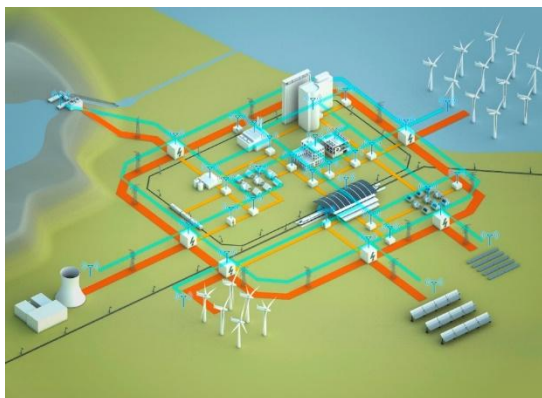


Fig. 2 Smart grid

#### V. THE ANALYSIS OF SITUATION IN A BUILDING AND ITS SURROUNDINGS

Several sensors were located in the building and its surroundings that are to measure a temperature of environment. They measure temperature in specific rooms and in the exterior. In the exterior are located also motion sensors. Two types of motion sensors are used. The first type covers sensors belonging to alarm device Paradox that are located in all rooms and on a terrace. The second type are exterior motion sensors that do not belong to Paradox and are located on strategically advantageous places in that way that they cover all parts of the area. Sensors of humidity and sensors of speed and direction of wind help to determine actual weather.

Other sources of information are devices using SNMP protocol – these devices send information about their condition to SQL database. Many of devices are monitored by means of relay that is able to communicate by Spinel protocol from Modbus, e.g. condition of lights (if they are switched on or off), particular parts of a circuit etc.

#### VI. PROPOSAL OF SOLUTIONS FOR SPECIFIC CONDITIONS

In the process of solutions proposal were given initial demands based on which was created a design depicted in the Figure 3. Devices that were available were utilized, as a suitable part of whole solution, in order to decrease the costs connected with implementation.

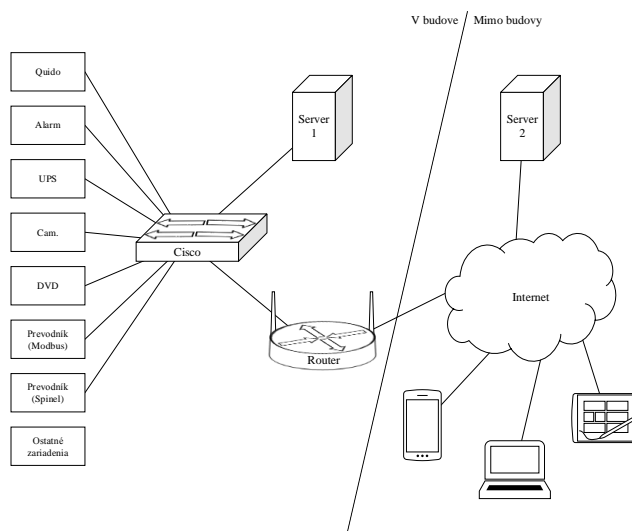


Fig. 3 View of all devices

First of all a list of all devices that will connect to the system was created, followed by

creation of a scheme that describes all the interconnections. A change-over switch Cisco Catalyst 2950 was available for connecting all other devices. The device Quido serves to monitor and set particular devices that are connected to it, connected to it are also cameras, alarm systems, UPS, DVD player and devices that communicate through protocols Modbus and Spinel. Other devices that connect to switch are Server 1 and router. Server 2 and all other devices that are able to display information about condition of devices as mobile phones, tablets, and notebooks are connected to the Internet.

In order to enable realization of the project, it has been necessary to update the wiring in the building since the building was built at the beginning of 1970s – aluminous leads, old circuit breakers and other wiring elements that cannot follow modern security standards were used nor they do enable the required control features. During the update of wiring new electric leads, other security elements such as overvoltage security and other elements that enable to control and monitor the building were installed. All the information that is being gained is further processed and saved in PostgreSQL database. It contains information as received by motion sensors, thermometers, humidity meters, anemometers and other devices from which there is a need to know their functions and condition.

#### A. PROGRAM CHATA APP

Program Chata App has been developed for observation of conditions of particular devices. This program is developed on platform JavaFX – this platform is built on the base of Java and simplifies programming GUI. Thanks to this platform is the program executable as web application or classic desktop application on all ordinary devices where JRE (Java Runtime Environment) is installed.

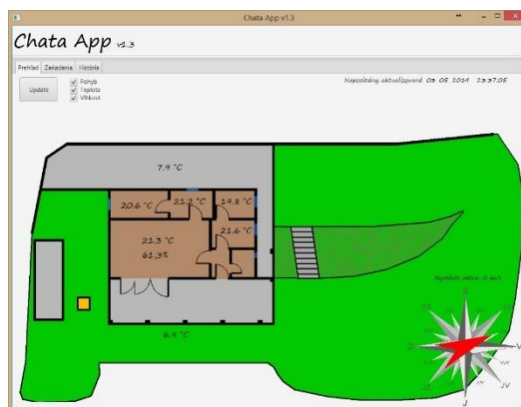


Fig. 4 Chata App – Overview

In Figure 4 is depicted plan view and cottage with rooms. On this map we may immediately determine actual temperature as well as humidity. It is also possible to watch movements on the plan view. This movement is shown as red colored sector on which was the movement detected. In the right lower corner there is information about actual direction and speed of the wind. Compass card shows cardinal points and red arrow show the direction of wind. Particular layers are able to be shown at once or separately. On website are presented actual information that are automatically updated. It is also possible to launch this update manually with button Update.

Overview of devices is shown in the Figure 5. In this bookmark user is able to find out condition of chosen devices.



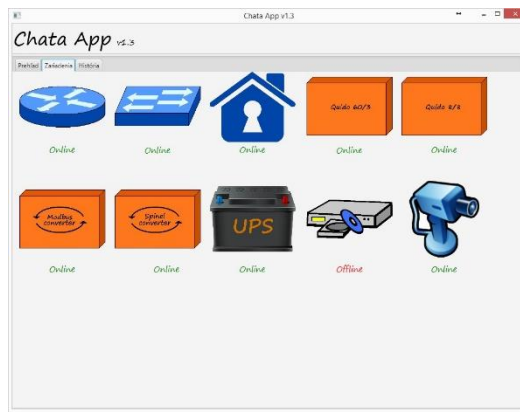


Fig. 5 Devices

Simple browsing of temperatures on different places and in different times is enabled. The consumer is able to choose specific season from year, month or day. It means that he can see the history of temperatures in hours, days or months. Information may be in a form of the chart or shown as a diagram, screenshot of just described is shown in the Figure 6.

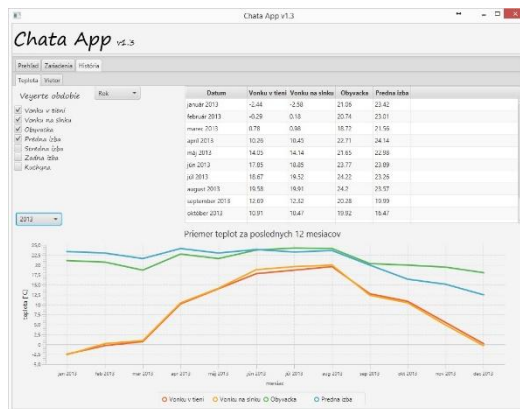


Fig. 4 History

## VII. POSSIBILITIES AND CONTRIBUTIONS OF THE SYSTEM

This system enables connection of huge amount of devices from different producers and subsequently managing of these devices on the bases of different incidents. The system is able to monitor actual condition in the building according to the program and it is able to control critical events based on given input. In present days this problem adjusts to concrete demands in concrete building but in the future it might be improved in the way that it would be relatively easily portable. During the process of the system development all the given requirements were successfully met.

The main contribution presented in this paper is the possibility to monitor conditions of devices and weather in the cottage from the distance as well as automatic control of some parts as lights, water consumption or heating.

## VIII. CONCLUSION

In this paper are summarized some of the commercially available solutions of intelligent houses and their status on Slovak and Czech market. Possibilities of smart meters and their use in real environment was discussed. Attention was also focused on intelligent networks Smart Grids that are being built nowadays and have a great potential in delivery of different energies to the customers.

We described their advantages and disadvantages and summarized their reliability and security. The aim, during the system proposal, was to use all available sources and minimize

costs of implementation. System is universal and it is able to connect devices from different producers and monitor and control their conditions. The system is able to gain information from sensors about conditions of devices and based on this information control them.

This system is to be further developed to enable adding of new devices to the system. Some of the devices are not able to connect because they are dependent on external suppliers, others have been connected for a trial.

#### ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] Derek Clements – Croome : Intelligent buildings: Design, management and operation ,Thomas Telford Limited, 2004 3-7s. ISBN 0-2772-3266-8
- [2] PRŮCHA, Jan: Není inteligentní dům jako inteligentní dům In: Inteligentní budovy č. 9 (2012), s. 8-9
- [3] PONCAROVÁ Jana: Poptávka po chytrých instalacích neustále roste In: Inteligentní budovy. č. 3 (2013), s. 31-32
- [4] ŠTEVO, Stanislav: Smart Meter – nejlepší zarobené peníze sú ušetrené peníze In: IDB JOURNAL č. 2 (2013), s. 19-22
- [5] Sinopoli, James: Signály a systavy : Systems for Architects, Owners, and Builders Elsevier: 2010. 65-65s. ISBN 978-1-85617-653-8
- [6] Meririan : Reading my meter. [Online] Dostupné [22.10.2013] <http://www.meridianenergy.co.nz/my-account/my-bill/reading-my-meter/>
- [7] Parliamentary Commissioner for the Environment: Smart electricity meters: How households and the environment can benefit. [Online] Dostupné [1.10.2013] <http://www.pce.parliament.nz/publications/all-publications/smart-electricity-meters-how-households-and-the-environment-can-benefit/>
- [8] eFOCUS: Smart grid/smart metering – trendy v zavádzaní inteligentných meracích systémov v utilitách. Očakávaná verzus realita na Slovensku. [Online] Dostupné [22.11.2013] <http://www.efocus.sk/konferencie/detail/smart-grid-smart-metering-trendy-v-zavadzani-inteligentnych-meracich-system/>
- [9] Amin, S.M. - Wollenberg, B.F. : Toward a smart grid: power delivery for the 21st century In: Power and Energy Magazine, IEEE č. 5 (2005), s. 34-41

# Web application of smart home

<sup>1</sup>Matej HUSÁR, <sup>2</sup>Michal KOVALČÍK, <sup>3</sup>Peter FECILAK, <sup>4</sup>František JAKAB

<sup>1,2,3,4</sup>Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

<sup>1</sup> matej.husar.2@student.tuke.sk, <sup>2</sup> michal.kovalcik@tuke.sk, <sup>3</sup> peter.fecilak@tuke.sk,  
<sup>4</sup> frantisek.jakab@tuke.sk

**Abstract** — This paper deals with the topic of intelligent systems. Web application for controlling household components is designed and implemented. Final solution is capable of communicating with actual hardware equipment by means of own communication protocol. Automatic generator of source initialization code for Arduino and Android application is implemented. The paper compares modern commercial systems for intelligent houses, along with connection methods and a consideration of advantages and disadvantages is presented.

**Keywords** — Smart home, Arduino, FuelPHP, HTTP communication

## I. INTRODUCTION

Penetration of advanced technologies in the areas of human life is consequence of rapid increasing development technology. This phenomenon has a major impact on the system and technical security solutions feature modern homes, causing meet customer requirements for higher living comfort and enhance safety in conjunction with respect for the environment.

With simple and clear web application for control the smart house it is possible to provide access to basic functions of smart home directly from home or outside. With regard to user comfort it is necessary to ensure easy home automation control components of hardware devices that are most commonly used by end user.

## II. GOALS

Task is to design and create a web page with the presentation and administration interface. Presentation part will contain basic information about system that was created. Administration part will be a control of electrical appliances of intelligent household. Electrical appliances will be connected to the control unit of Arduino device by Ethernet Shield. Communication between Web applications and control equipment Arduino will be sent via HTTP requests. For the task, it is necessary to study in detail the technology of HTTP communication between control device and the web application. It is also necessary to invent own communication protocol through which the communication will be encrypted and only recognizable by Arduino device and web applications.

## III. COMMERCIAL SYSTEMS OF SMART HOME

### A. Central system inHome

InHome central system is a product of the Czech company called Insight Home, which is leader on the field of intelligent households. InHome system [1] is built on the basis of control elements created by the American Society AMX. As reported by [2], control system can also be supplied through the drivers of U.S. Company AMX. However these drivers are so complicated that people with disabilities might find it difficult to control them. System wiring system Insight Home is modular and easily expandable with new control modules. Individual modules can be extended gradually.

Benefits of this system include simple and intuitive control of appliance and wide spectrum of active control devices. Disadvantages of system are complex drivers and inability of developers to create their own applications for the system inHome.

#### *B. Smart home system Control 4*

The company Control 4 offers a complete system for controlling households from simple management of audio devices to complex controlled heat ventilation. Authors in [3] state that user interface is redesigned so users can use the system for mobile devices, tablets, laptops and even desktop computers.

The big advantage of Control 4 is a wide variety of controls, this control is aimed also for disabled people. Another advantage Control 4 system, according to [3] its own application store, designed specifically for smart home with custom application development. The disadvantages belong only to highly initial investment.

#### *C. LG system Smart Home Service*

These LG smart home technologies are used directly in the manufacture of certain products from LG. Every product containing these technologies provides the possibility for a user to control appliances or electronics in the home [4].

According to the source [4], each appliance is equipped with SmartControl and SmartShare, and this appliance can be also controlled via the SmartControl, which is freely available in store LG. The application is also available for the operating system Android and iOS. Control of appliances is also possible via television from LG that has a pre-installed application.

#### *D. Smart home system Loxone*

Intelligent Loxone is used for home automation and provides a wide range of management appliances from simple sunblind control to smart management of heating regulation [5]. This intelligent system was designed by the Austrian company. Loxone is based on mini server Loxone, this server forms the basis of home automation and controls all devices in household. Mini server provides many digital inputs and outputs, whether in digital or analog form.

Mini server, according to [6], is equipped with a simple Loxone Config configuration utility. This utility allows easy to adjustment of the lighting, automatic control of room temperature, shows visualization appliances in the room and then saves the changes to mini server. According to the [6] other modern technology provided by the company Loxone is called Loxone Air technology – it was developed for easy operation of wireless smart homes.

### IV. SYSTEM DESIGN

The Web system is to provide users with administration and presentation interface through which they can control devices in an intelligent household in real-time. Important device utilized in whole project is Arduino. All requests that are going from the client are processed by the web application server using MySQL database server. Each device of the scheme must be actively connected to the local network.

#### *A. Communication protocol*

Communication between web application and Arduino microcontroller is based on HTTP requests. Every request includes unique string which is sending in URL address like parameters with name of command.

Communication protocol consists of following:

- CD [ number of household] - unique identification of household.
- CM [ number of rooms] - unique identification of room.
- CT [ type of appliance] - identification of appliance type.
- CS [ number of appliance] - unique identification of appliance.
- H [ value] – value of appliance.

This format ensures unique identification of household and uniform string for communication with Android and Arduino device.

#### *B. Presentation interface*

Presentation interface provides basic information, i.e. following:

Basic information about product of smart home with contact form.

Articles about new features of product and form for login into administration section.

#### *C. Administration interface Presentation interface*

Administration interface includes all control panels for management of household and users. This interface provides an active control of intelligent household elements.

Abstract illustration of the system design is shown in the Figure 1.

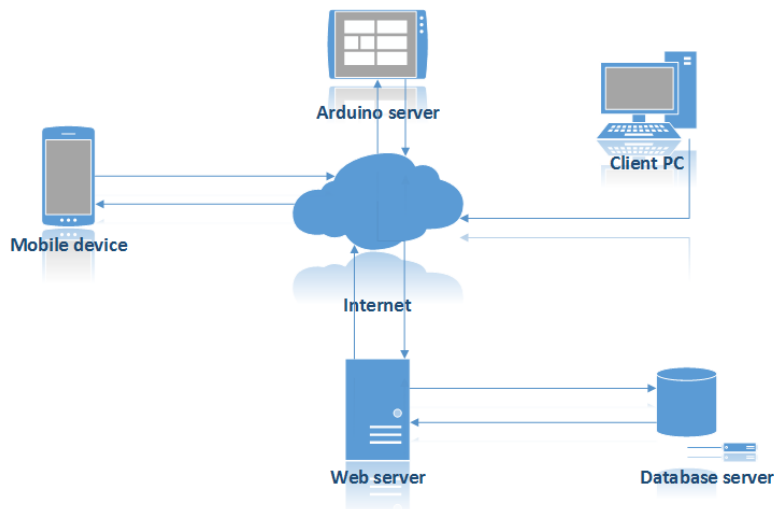


Fig. 1 Abstract illustration of smart home

## V. AUTOMATICALLY GENERATOR DECLARATION AND INITIALIZATION CODE

Administrator of web application has ability to create configurations of household. These configurations are possible to generate into format which can be easily process by Arduino or Android application.

### A. Generator for Arduino device

Format of generated, declared and initialized string is well defined. This format provides easy implementation of the appliances in the Arduino device. Now is needed just to copy the generated string, burn it into Arduino device memory and set the desired addresses for thermometers and buttons. Generated format contains all the necessary information about the appliances. In the Figure 2 there are displayed generated rooms with the initialization string for the Arduino device.

```
Appliance appliances [1]
Button buttons [8]
appliances[1] = Appliance(10, 17, 1, 31);
buttons[8] = Button(,);
Appliance appliances [1]
Button buttons [8]
appliances[1] = Appliance(10, 17, 3, 32);
byte SN temp*8[] = {};
appliances[1] = Appliance(10, 17, 3, "1");
buttons[8] = Button(,);
```

Fig. 2 Example of generated declaration and initialization string for Arduino

This initialization and declaration string defines number of appliances, numbers of button in rooms, definitions of thermometers and thermostats. Also definition mentioned above clearly defines house, room and appliance unique identification.

### B. Generator for Android application

Format of generated, declared and initialized string for Android application is not precisely defined. Format has particularly informative value and serves as a configuration model for mobile applications. User on the basis of this model will create a copy of any already established household configurations.

## VI. IMPLEMENTATION OF WEB APPLICATION

### A. Presentation interface

The entire presentation section is enveloped in a graphical Bootstrap framework. Framework, in conjunction with the appropriate use of HTML and CSS code, causes the page design to be correctly and legibly displayed on mobile devices, tablets and high-resolution monitors. Based on your monitor resolution, the graphic elements will be shown differently on the web application, so it will be easy to navigate. Sample of the change of graphic elements is shown in the Figure 3.



Fig. 3 Example of presentation interface on mobile device and high resolution monitor

### B. Administration interface

Administration interface of web application is used for the management of users and households. From this interface it is possible to control active elements of intelligent household. Controls of household are drawn on the basis of how they were defined by administrator. From all entered data is rendered left rail, which contains clickable household rooms. When clicked on room, control elements of appliances are plotted and the request is sent on Arduino device from which are current values of indicators appliances detected and real values determined from Arduino device are set. After drawing controls, the user can control appliances. At each change the request is sent in a predefined format for the Arduino device and the changes are saved on database server.

## VII. CONCLUSION

The initial point when solving the issue of intelligent household was to analyze the existing intelligent systems. Given the amount of existing systems, the comparison was focused on several factors, e.g. connection method, description of system modules or affordability. The main task of this paper was to propose a design and implement a web application through which we will control appliances connected to the Arduino microcontroller.

The next task was to design and implement an automatic generator initializing source code for the Arduino device via web application and application on the Android platform. This initialization of source code for the Arduino control device has a clearly defined structure which makes it possible to easily implement the configuration. Initialization code ensures easy portability of configurations. Generator for Android was designed only as an export to desired format and has only informational value.

Communication protocol for communication between Web applications and control device Arduino was designed and implemented. This communication protocol precisely defines the configuration information in household and provides a space for future expansion of new types of appliances. Arduino device is reliable and powerful enough to meet the role of managing smart home devices.

#### ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] INSIGHT HOME: Řídicí systém inHome AMX. [2013-10-28] Available at: <<http://www.insighthome.eu/inHome.html#Home>>.
- [2] INSIGHT HOME a.s.: Systémová elektroinštalácia – Základ inteligentnej a bezpečnej budovy. [2013-12-22] Available at: <[http://odbornecasopisy.cz/data-ftp-user/PRAGOSMART\\_2012/Insight-Home\\_Pragosmart\\_2012.pdf](http://odbornecasopisy.cz/data-ftp-user/PRAGOSMART_2012/Insight-Home_Pragosmart_2012.pdf)>
- [3] CONTROL 4: Basics. [2013-11-1] Available at: <<http://www.control4.com/basics>>.
- [4] CHOVAŇÁK, Fero: Inteligentná domácnosť so spotrebičmi od LG. [2013-11-2] Available at: <<http://www.mojandroid.sk/2013/01/11/inteligentna-domacnost-lg-ces-2013>>.
- [5] ŽEMLIČKA, Petr: Přes 70 nových funkcí! Loxone Config 5 – Nejlepší software pro chytré domy všech dob! [2014-4-13]. Available at: <<http://blog.loxone.com/cscz/pres-70-novych-funkci-loxone-config-5-nejlepsi-software-pro-chytre-domy-vsech-dob>>.
- [6] ŽEMLIČKA, Petr: Přestavba na inteligentní dům nebyla nikdy jednodušší. [2014-4-13] Available at: <<http://blog.loxone.com/cscz/nove-prestavba-na-inteligentni-dum-nebyla-nikdy-jednodussi>>.

# Control module for smart house based on Arduino and its modules

<sup>1</sup>Marek KURILEC, <sup>2</sup>Michal KOVALČÍK, <sup>4</sup>Peter FECILAK, <sup>4</sup>František JAKAB

<sup>1,2,3,4</sup>Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

<sup>1</sup>marek.kurilec.2@student.tuke.sk, <sup>2</sup>michal.kovalcik@tuke.sk, <sup>3</sup>peter.fecilak@tuke.sk,  
<sup>4</sup>frantisek.jakab@tuke.sk

**Abstract** — This paper analyzes suitability of utilizing 1-Wire technology for the development of sensor network for smart houses and further presents designs of a solution for Arduino application. Arduino board is used as a control module of smart house. Such control module is controlled via mobile device running Android or using web application. Control module was implemented into home environment. Conclusion summarizes advantages and disadvantages of designed module.

**Keywords** — 1-Wire technology, Android, Arduino, communication protocol, smart home, web application

## I. INTRODUCTION

The main idea of this project is to utilize freely available and inexpensive solutions. The paper is concerned with the issue of smart homes – more specifically, the development of a control module for smart home using the Arduino microcontroller and its modules. The goal is to design an application for Arduino to manage smart home control from mobile devices running Android and web application. Final solution utilizes open-source Arduino Uno microcontroller and its extension module Ethernet shield. To realize communication between microcontroller and mobile device or web application is created unique communication protocol based on HTTP GET requests and JSONP method.

## II. 1-WIRE TECHNOLOGY AND ITS USE

### A. Technology description

Dallas Semiconductor has designed and developed technology based on a single bus master and multiple slaves that transmits digital communications and operating power for the slaves over a single twisted pair cable. An important factor of the technology is that every slave has a globally unique digital address. 1-Wire uses a single wire (plus GND) to accomplish both communication and power transmission, the technology has found application in a number of areas.

The 1-Wire net is a low cost bus based on a PC or microcontroller communicating digitally over twisted pair cable with 1-Wire components. The network is defined with an open drain master/slave architecture that uses a resistor pull-up to a nominal 5V supply at the master. A 1-Wire net-based system consists of three main elements: a bus master with controlling software, wiring and associated connectors and 1-Wire devices. The system allows tight control because no node is allowed to speak unless requested by the master, and no communication is allowed between slaves, except through the master. Within each 1-Wire slave created is stored a ROM section with its own guaranteed unique, 64-bit serial number that acts as its node address [1][2][3].

### B. Suitability for smart house

The development of 1-Wire technology made possible the combination of electronic communication and instrumentation based on positive identification of individual nodes on a single self-powered net. Continued growth of the technology has produced an increasing array of 1-Wire chips that were able to provide an interface with the environment measuring events, voltage, current, temperature, position, etc. These chips enabled the construction of sensors



that measure a host of environmental parameters on a single twisted pair. Temperatures, amount of sunlight, barometric pressure or humidity are easily measured with 1-Wire sensors over 1-Wire net. These are all factors that affecting personal comfort.

### III. CONTROL MODULE FOR SMART HOUSE

The control module consists of hardware and software part. The hardware part is a microcontroller Arduino Uno with Ethernet shield plugged into Wi-Fi router placed in the smart house. The software part is an application which is uploaded to microcontroller's memory.

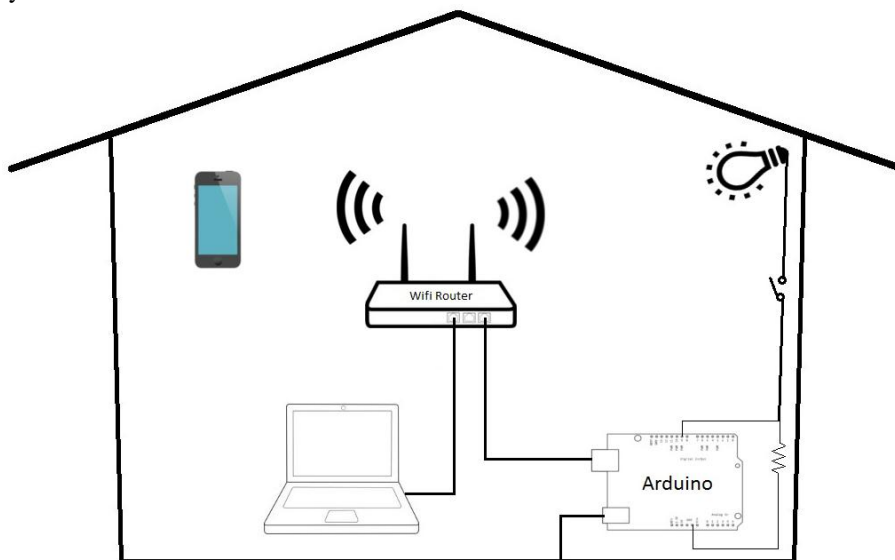


Fig. 1 Control module Arduino plugged into local Wi-Fi router

#### A. Hardware – Arduino Uno and Ethernet shield

The Arduino Uno is a microcontroller board based on the ATmega328. Arduino is an open-source, prototyping platform. The board has 14 digital input/output pins of which 6 can be used as PWM outputs and 6 analog inputs (larger models provide more of those input/output options). Further it has a 16 MHz crystal oscillator, 32kB of flash memory, a USB connection, a power jack, an ICSP header, and a reset button. Part of the Arduino package is a programming environment, where code is written in simplified C-like language, and transferred to the Arduino using a USB cable. After programming, an Arduino can work while being connected to a PC or operate standalone [4].

The Ethernet Shield allows an Arduino board to connect to the internet. It is based on the Wiznet W5100 Ethernet chip. The Wiznet W5100 provides a network (IP) stack capable of both TCP and UDP. It supports up to four simultaneous socket connections. The Ethernet shield connects to an Arduino board using long wire-wrap headers which extend through the shield. This keeps the pin layout intact and allows another shield to be stacked on top [5].



Fig. 2 Arduino Uno with plugged Ethernet shield

*B. Software – main application for Arduino and communication protocol*

Communication protocol is defined as a unique string that represents a request send by the user. User as a client sends a request to the server - Ethernet shield. Unique string is used to set values for appliances, e.g. adjust the amount of light in the living room. This is an example how the unique string looks like: “*CD1:CM3:CT2:CS9:H255*”, where:

- **CD#** - represents ID of the house. One user can have more than one house.
- **CM#** - represents ID of the room in which is placed particular appliance
- **CT#** -represents ID of the appliances. Each appliance in house has own unique ID.
- **CS#** - represents ID of the particular appliance, e.g. in one of the room can be placed more than one independent light so all must have own ID.
- **H#** - represents value which is sets on particular appliance.

After sending request, the client receives a response from the server. JSONP communication technology, that allows web application requests server to send data, is utilized. The web application uses the GET method which joins data from the client to the servers URL. The resulting URL has a unique syntax which the client must always keep, e.g. URL containing data: “*http://192.168.1.10/?format=JSONP&command=CD1:CM3:CT2:CS9:H255&callback=X*”. Only valid requests are handled, while invalid request is not performed and the client does not receive a response. Once server received valid request, it then sends message back to client in JSONP format and as well sends standard HTTP response header. This header indicates that request was received and handled. Client parses the message then sets appliances to the actual value and displays it.

The main application consists of three sections. The first section is declaratory with a list of imported libraries. Followed by the declaration of 1-Wire bus, i.e. with which pin number is 1-Wire bus associated. Next MAC address and IP address is declared. It is necessary to define MAC address manually since the module does not one assigned by the manufacturer. IP address can be assigned statically or dynamically – dynamic is using function that asks the DHCP server for IP address assignment, while in our case is used a static IP address assignment. The program is designed to be the most generic. Some parts of the source code for the Arduino application are generated by web applications. The generated code is added to the application and loaded into memory of the board. The next section of the application consists of a function called `setup()`. This function is called only once after switching on the board or by pressing the reset button on the Arduino board, most of the part consists of generated code. Last section forms a function `loop()` which is carried out in cyclic loop. In this function is performed checking of received requests. Based on section of request that starts with “*command=*” is performed specific function. Program sends a message to a client in JSONP format with states of appliances – received request is “*command=STATE*”, or performs the function that sets state on particular appliance – e.g. received request is “*command=CD1:CM3:CT2:CS9:H150*”. As a part of function `loop()` are created functions that set lights, check current temperature or set thermostat.

IV. IMPLEMENTATION OF CONTROL MODULE

Final control module was placed into three-room apartment. Deployment of appliances in each room is show in following table.

Table 1  
 Deployment of appliances in each room of apartment

Room	Appliances
Hall	2 light circuits
Kitchen	2 light circuits, 1 thermometer
Kids room	1 light circuit, 1 thermometer
Bedroom	1 light circuit, 1 thermometer
Living room	1 controllable light, 1 thermometer, 1 thermostat
Bathroom	2 light circuits, 1 thermometer
Toilet	1 light circuit

Applied system consists of the following components:

- control module Arduino Uno,
- Wi-Fi router,
- data line,
- power cable,
- switch unit,
- dimmer.

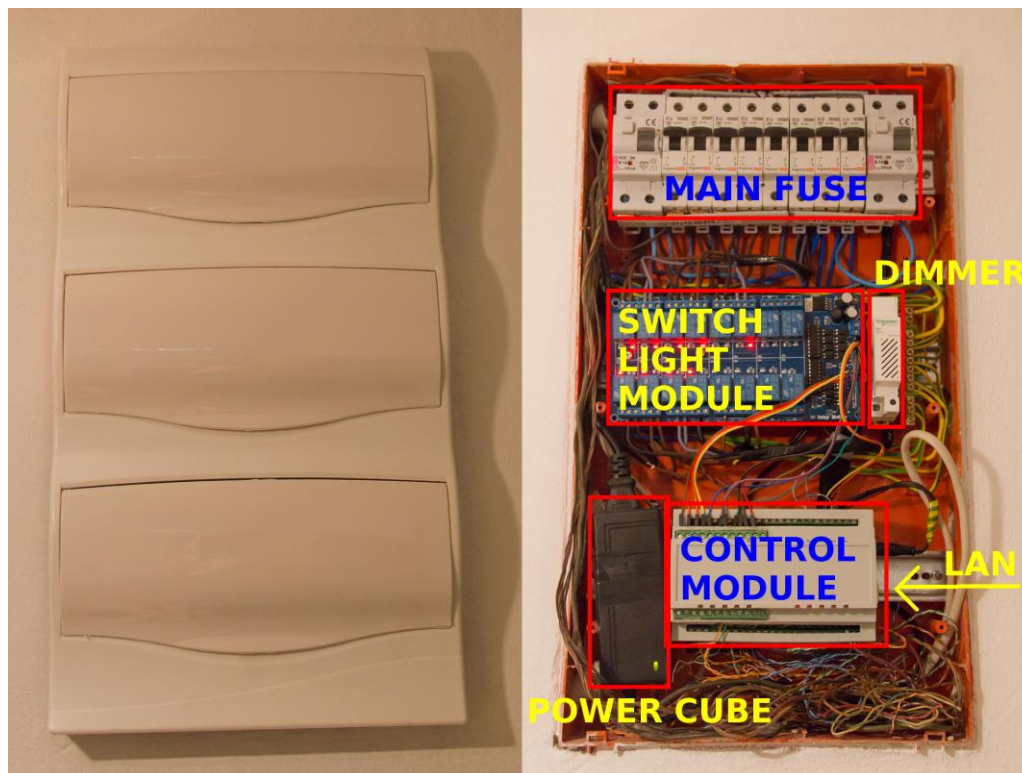


Fig. 3 Wiring of smart system into the house

Web application generated necessary parts of the code that were added to source code of Arduino application and subsequently were loaded into the memory. We were able to change lighting in home and read temperature in rooms using of the local Wi-Fi network and mobile device connected to it. Values of lighting were changing smoothly and without delay. Whole system worked correctly.

## V. CONCLUSION

In this paper was presented implementation of appropriate technology for smart home environment. Design of the control module based on Arduino board and its modules that are controlled via mobile device running Android or though the web application was described in detail.

Paper concludes that the management of smart home can be implemented using various microcontrollers and additional modules. Microcontroller Arduino Uno is an appropriate solution utilized as a control module for smart houses. Major advantage of this system is affordability, programmability and expandability using additional modules. As a restriction may be considered low number of I/O pins on the board Arduino Uno. The solution of this limitation might be replacement of microcontroller Arduino Uno for Arduino Mega which has a larger number of I/O pins however Mega is more expensive version. Developed application provides the elementary functions that are further expandable and adaptable to user's needs.

## ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] B. Linke, *Overview of I-Wire technology and its use*, 2008. Available at: <http://www.maximintegrated.com/en/app-notes/index.mvp/id/1796>
- [2] D. Penezic, *I-Wire communication interface*, 2010. Available at: <http://www.instructables.com/id/1-wire-communication-interface/>
- [3] S. Najar, *Simulation of I-Wire sensors*, Stockholm, 2012. Available at: <http://kth.diva-portal.org/smash/get/diva2:563082/FULLTEXT01.pdf>
- [4] *Arduino Uno – Overview*. Available at: <http://arduino.cc/en/Main/arduinoBoardUno>
- [5] *Arduino Ethernet Shield – Overview*. Available at: <http://arduino.cc/en/Main/ArduinoEthernetShield>

# Aplikácia slovníkového prístupu pri analýze sentimentu

*Peter KOVÁČ, Kristína MACHOVÁ, Martina TARHANIČOVÁ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

peter.kovac.10@student.tuke.sk, kristina.machova@tuke.sk,  
martina.tarhanicova@student.tuke.sk,

**Abstrakt** — V dnešnom svete dokážu človeka ovplyvniť názory niekoho iného a teda ovplyvniť jeho ľudské chovanie. Naše presvedčenie voľby, ktoré konáme okolo seba, všetko je to odrazom informácií, ktorými sme ovplyvňovaní a ktoré získavame prostredníctvom dnešných výdobytkov modernej doby. Najväčším poskytovateľom informácií je internet. Poskytuje nám veľké množstvo recenzií, názorov ľudí na produkty či služby a mnoho ďalších informácií. Všetky tieto informácie je možné spracovávať a analyzovať pomocou analýzy sentimentu, ktorá sa v poslednej dobe dostáva do popredia. Je odbor, ktorý analyzuje názory ľudí, ich pocity, hodnotenia, posudky, postoje a emócie voči subjektom, ako sú produkty, služby, organizácie, problémy, udalosti, témy a ich atribúty. Je jednou z najviac aktívnych oblastí v rámci spracovania prirodzeného jazyka. Cieľom práce je podať stručný a výstižný pohľad na analýzu sentimentu, objasniť význam analýzy sentimentu, jeho využitie pre ľudí a prínosy pre dnešnú spoločnosť. V práci sú popísané metódy klasifikácie, ktoré využívajú analýzu sentimentu. Hlavným cieľom práce bolo zistiť emócie vyjadrené v slovenských dokumentoch pomocou slovníkového prístupu, ktorý využíva nami vytvorený klasifikačný slovník emocionálnych slov. Posledným bodom práce je vypočítať a vyhodnotiť výsledky presnosti a návratnosti klasifikovaných dokumentov.

*Kľúčové slová* - analýza sentimentu, klasifikácia sentimentu, dolovanie názorov, n-gram

## I. ÚVOD

Analýza sentimentu je čoraz viac využívanou metódou spracovania prirodzeného jazyka. Je to kvôli väčšiemu využívaniu sociálnych sietí ako je Facebook alebo Twitter. Vďaka sociálnym sieťam, blogom, fóram alebo recenziám môžeme z textov tzv. dolovať dáta a zisťovať emócie, postoje alebo názory ľudí na subjekt. Využitie analýzy sentimentu pre získavanie názoru ľudí je v dnešnej dobe veľmi potrebné najmä v priemysle a obchodnej sfére, pretože názory iných ľudí majú veľký vplyv na ich chovanie. Firmy vďaka analýze sentimentu môžu získavať spätné informácie od zákazníkov o produktoch a tým prispôsobiť jednotlivé subjekty pre zákazníka a zvýšiť predajnosť, konkurenciu a schopnosť reagovať na požiadavky zákazníkov. Ďalšou využiteľnou sférou pre analýzu názorov je politická scéna, kde sú názory voličov dôležitým faktorom pre existenciu politických strán. Na základe týchto názorov môžu zlepšovať svoj volebný program a tak dosiahnuť lepšie výsledky.

Väčšina prác zaoberajúcich sa analýzou sentimentu sa zaoberá hodnotením anglických dokumentov. Dôvodom je že angličtina je celosvetovým jazykom a najviac príspevkov na diskusných fórach, či sociálnych sieťach sa nachádza práve v angličtine, čím je zaručená veľká množina vstupných dát. Preto aj v práci sme väčšinou využívali anglické odborné články, kde sa zaoberali analýzou sentimentu, jej využitím a problematikou. Keďže analýza sentimentu stále rozvíja svoje využitie a na Slovensku je málo aplikácií ktoré by pracovali s naším národným jazykom mojím cieľom bolo. Vytvoriť aplikáciu, ktorá by používala slovenský jazyk a na základe jeho slov by hodnotila emócie vyjadrované v slovenskom jazyku.

Práca je delená na päť tematických častí, ktorým sa postupne vyjadríme. Prvá časť práce sa zaoberá podaním stručného obsahu o histórii sentimentu, vysvetlením čím sa zaoberá analýza sentimentu a popisom jej oblastí využitia.

V druhej časti sme sa venovali popisu metódam kontrolovaného a nekontrolovaného učenia, jej cieľom je popísať existujúce metódy analýzy sentimentu, ich problematiku a využitie.

Tretia kapitola je venovaná hlavnému cieľu tejto práce, návrhu vhodného slovníkového prístupu pre klasifikáciu dokumentov a vytvoreniu slovenského emocionálneho slovníka a ich implementácie, ktorá je popísaná v ďalšej časti práce. Pre vhodnú klasifikáciu dokumentov sme použili slovníkový prístup, založený na porovnávaní slov. Porovnávanie slov klasifikovaného

dokumentu sa vykonávalo pomocou nami vytvoreného slovníka, obsahujúceho slovenské slová nesúce jednu z emócií (odpor, hnev, radosť, smútok, strach, úžas). Nachádza sa tu stručný popis obsahu vytvoreného slovníka a jeho štruktúra.

V Nasledujúcej štvrtej časti je popísaná implementácia navrhnutého slovníkového prístupu pomocou algoritmov uvedených v **Chyba! Nenašiel sa žiaden zdroj odkazov..** kapitole. V tejto práci sme sa venovali hodnoteniu anglických a slovenských dokumentov. Pričom anglické boli hodnotené pomocou slovníka SentiWordNet a výstupom bol pozitívny alebo negatívny dokument. Slovenské hodnotené dokumenty využívajú už spomenutý slovník emocionálnych slov a ich výstupom je najviac vyjadrená emócia (napr. : odpor, hnev, radosť, smútok, strach, úžas) v dokumente.

Poslednou piatou časťou sú experimenty, výpočet potrebných hodnôt pre vyhodnotenie presnosti a návratnosti jednotlivých emócií nachádzajúcich sa v klasifikovaných dokumentoch.

## II. SENTIMENT

### A. Analýza sentimentu

Odbornú definíciu analýzy sentimentu uviedol Bing Liu v knihe *Sentiment Analysis and Opinion Mining* : „ *Analýza sentimentu, často nazývaná ako dolovanie názorov (opinion mining), je vedný odbor, ktorý analyzuje ľudské názory, pocity, hodnotenie, postoje a emócie voči subjektom, ako sú výrobky, služby, organizácie, jednotlivci, problémy, udalosti, témy a ich atribúty* “ [1].

Analýza sentimentu využíva spracovanie prirodzeného jazyka (NLP) a analýzu textov. Preto výskum v analýze sentimentu má nielen významný vplyv na NLP, ale môže mať dostatočný vplyv aj na vedu, politiku, ekonomiku a sociálne vedy, pretože sú ovplyvnené názormi ľudí.

Lubovoľný text je možné rozdeliť podľa toho či obsahuje alebo neobsahuje sentiment a to dvomi spôsobmi :

- objektivne – neobsahuje sentiment
- subjektívne – obsahuje sentiment

Tento problém je možné riešiť pomocou algoritmov strojového rozpoznávania (napr. AdaBoost alebo BoosTexter). Ide o učiaci sa algoritmus, ktorému je najprv treba dodať vstupné dáta, pomocou ktorých je schopný robiť binárnu klasifikáciu, teda rozhodovať sa do ktorej množiny zadaný text patrí [2].

Sentiment môžeme skúmať v článkoch, vo vetách, recenziách, hodnotení , na sociálnych sieťach, v blogoch a fórach vyjadrujúcich niekoho názor, alebo aj u jednoduchých fráz. To umožňuje nielen nájsť slová, ktoré vyjadrujú sentiment, ale nájsť aj vzťahy medzi viacerými slovami vyjadrujúcimi sentiment a presne ich identifikovať. Dokumenty alebo slová môžu z hľadiska sentimentu vyjadrovať [3]:

- pozitívni sentiment
- negatívni sentiment
- neutrálny sentiment

### B. Využitie analýzy sentimentu

V dnešnej dobe veľké množstvo spotrebiteľsky generovaných médií (CGM), ako sú fóra, hodnotenia či novinové články prinášajú skvelé príležitosti, ale aj riziká pre dnešné firmy. Najmä v automobilovom priemysle, elektronike alebo filmovom odvetví, kde zákazníci písali recenzie o výrobkoch alebo ich vlastnostiach. Zberom a analýzou týchto hodnotení by mohli noví zákazníci nájsť názory druhých na vlastnosti a kvalitu výrobku. Takto môžu porovnávať výrobky aby našli ten najlepší, ktorý vyhovuje ich potrebám. Firmy by mohli svoju kartu stavať na obsah CGM, čím by pochopili názory spotrebiteľov o ich produktoch a službách, zistili by si silné a slabé stránky svojich produktov a tým by mohli vytvárať nové inovatívne príležitosti a výhody oproti konkurencii. Na druhej strane nesledovanie takýchto CGM by firme mohlo priniesť riziko v imidži a ich reputácii, pretože dnes je cez internet šírené veľké množstvo CGM. **Chyba! Nenašiel sa žiaden zdroj odkazov..**

Sociálne médiá obsahujú veľké množstvo dát bohatých pre analýzu sentimentu vo forme blogov, aktualizácií statusov alebo správ či odkazov medzi používateľmi. Sentiment analýza takýchto užívateľov sociálnych sietí je veľmi užitočná v dôsledku poznania názoru spoločnosti. Twitter analýza sentimentu je v porovnaní so všeobecnou analýzou sentimentu ťažšia kvôli

prítomnosti slangových slov a preklepov. Príspevky na sociálnej sieti Twitter môžu mať maximálne 140 znakov, ale často sú kratšie a obsahujú hypertextové odkazy, ktoré sú pre detekciu sentimentu irelevantné. Pre používateľa je analýza týchto textov príliš zdlhavá, preto sú na to použité rôzne techniky analýzy sentimentu. [4].

### III. METÓDY ANALÝZY SENTIMENTU

Existujú dve hlavné metódy pre klasifikáciu sentimentu:

- symbolické metódy
- metódy strojového učenia

Symbolické metódy vyžadujú rozsiahlu databázu vopred definovaných emócií a efektívnu reprezentáciu znalostí pre identifikáciu sentimentu. Prístup strojové učenie využíva rozvíjajúcu sa tréningovú množinu a sentiment klasifikátor, ktorý klasifikuje pocity. Vzhľadom k tomuto nie je potrebné pre strojové učenie mať vopred definovanú databázu emócií ako u symbolických metód, je to jednoduchšie [4].

### IV. SLOVNÍKOVÝ PRÍSTUP

Slovníkový prístup vysvetľuje, že textový dokument chápeme ako zoznam slov bez akýchkoľvek väzieb medzi sebou. Potom sa určuje sentiment každého slova a skombinuje sa s nejakou agregáčnou funkciou. Sentiment možno určiť pomocou viacerých metód s využitím externých zdrojov. Jedná sa o problém nekontrolovaného učenia a tento prístup využíva externé slovníky. Slovníky môžu byť zostavené manuálne, takýto postup je veľmi náročný a časovo zdlhavý, alebo môžu byť vytvorené automaticky pomocou tzv. seed words, ktoré slovníky automaticky rozširujú.

#### A. Návrh klasifikácie dokumentu

Pri klasifikácii textu sme v našej práci použili slovníkovo založený prístup, pričom sme využili nami vytvorený slovník emocionálnych slov. Prvým krokom k analýze názorov a emócií je prevod dokumentu na štruktúrované dáta. Teda prevod dokumentu na zoznam pričom prvkami zoznamu sú slová na základe ktorých sa určuje polarita, alebo emocionálne vyjadrenie daných dokumentov. Po získaní zoznamu slov je potrebné porovnanie tohto zoznamu so slovníkom, aby sme zistili ktoré emócie sú vyjadrené v dokumente. Pre získanie výsledku víťaznej emócie sú počítané pravdepodobnosti pre každú jednu emóciu (odpor, hnev, radosť, smútok, strach, úžas). Víťaznou emóciou, ktorá je vyjadrená v dokumente sa stáva tá, ktorá má najvyššie percentuálne ohodnotenie pravdepodobnosti. Posledným krokom pri klasifikácii dokumentu sú výpočty presnosti a návratnosti.

#### B. Slovník emocionálnych slov

Dôležitým bodom práce bolo vytvorenie slovníka emocionálnych slov v dostatočnom rozsahu. Na tejto práci sa podieľalo viacero študentov, ktorý tento vytvorený slovník využijú vo svojich prácach. Slovník zahŕňa slová rôznych slovných druhov, ktoré sú nositeľom významu:

- prídavné mená
- príslovky
- slovesá
- a podstatné mená.

Keďže slovenčina je krásny a jeden s najťažších jazykov, slová uvedené v tomto slovníku nenadobúdajú iba základný tvar, ale sú skloňované pre presnejšiu klasifikáciu dokumentov. Týmto prístupom sme dosiahli, že slovník obsahuje vyše 19-tisíc slov, ktoré sú usporiadané podľa abecedy. Tento slovník obsahuje pre každé jedno slovo číselné ohodnotenie, ktoré vyjadruje silu sentimentu slova z intervalu  $\langle -3;3 \rangle$ . Každému slovu je priradená emócia vystihujúca toto slovo. Emócie ktoré sa priradzovali k slovám sú:

- disgust
- angry
- joy
- sadness

- fear
- surprise

Tab. 1 Ukážka slovníku emocionálnych slov

BÚRLIVOM	ADJ	ANGRY	-2
BÚRLIVOU	ADJ	ANGRY	-2
BÚRLIVÚ	ADJ	ANGRY	-2
BÚRLIVÝ	ADJ	ANGRY	-2
BÚRLIVÝCH	ADJ	ANGRY	-2
BÚRLIVÝM	ADJ	ANGRY	-2
BÚRLIVÝMI	ADJ	ANGRY	-2
BÚRME	VERB	DISGUST	0

## V. IMPLEMENTÁCIA

### A. Anglické dokumenty

V našej práci používame aj slovník SentiWordNet 3.0 , ktorý je taktiež vytvorený na klasifikáciu dokumentov alebo dolovanie názorov (data mining). Tento slovník obsahuje zoznam ohodnotených slov z Wordnetu. Wordnet je slovníková databáza anglického jazyka, kde sú slová rozdelené podľa slovných druhov a slová sú navzájom prepojené sémantickými vzťahmi. V slovníku SentiWordNet je každému slovu priradená číselná hodnota, či je slovo pozitívne alebo negatívne. Číselná hodnota je z intervalu  $\langle 0,1 \rangle$ . Slovník je voľne dostupný na nekomerčné účely.

Klasifikácia anglického dokumentu pomocou slovníku SentiWordNet je postavená na základe pozitívnych alebo negatívnych hodnôt slova. So slovníka si potrebujeme uložiť do zoznamu anglické slovo a jeho pozitívnu a negatívnu číselnú hodnotu. Všetky slová nie sú v dobrom tvare, pretože za každým slovom nasleduje znak # a číslo. Pre porovnávanie slov je to zlé, preto musíme tieto znaky od každého slova odstrániť a tak ich uložiť do zoznamu. Ak už máme takýto zoznam slov na rade je porovnávanie jednotlivých slov z dokumentu so slovníkom. Ak sa slovo z dokumentu nájde v slovníku jeho pozitívna a negatívna hodnota sa pripočíta prislúchajúcim premenným a program pokračuje v hľadaní ďalej až pokým neprejde celý dokument. Po ukončení hľadania sa navzájom porovná pozitívna a negatívna hodnota a výsledok je buď pozitívny alebo negatívny dokument podľa toho ktorá hodnota je vyššia

### B. Slovenské dokumenty

Hlavným cieľom tejto práce je klasifikácia slovenských dokumentov, pre zistenie emocionálneho vyjadrenia dokumentu, pomocou nami vytvoreného slovníka emocionálnych slov. Pri tomto prístupe potrebujeme taktiež z dokumentu (testovaného textu/článku) vytvoriť zoznam slov. Pre vytvorenie zoznamu je potrebné z dokumentu odstránenie interpunkčných znamienok, prevedie veľkých písmen na malé a dokument rozdeliť na slová.

Po prevedení dokumentu na zoznam slov je potrebné s tohto zoznamu porovnať každý jeden prvok, teda slovo s emocionálnym slovníkom a zistiť, či sa nachádza dané slovo v slovníku. Ak je zhoda pozitívna, slovo sa našlo je ďalšou dôležitou úlohou uchovať si informácie o tomto slove. Hodnoty ktoré sa o slove uchovávajú sú:

- emócia prislúchajúca slovu
- číselné ohodnotenie sentimentu slova

Na zistenie týchto potrebných hodnôt sme v práci použili nami vytvorený emocionálny slovník v ktorom každému slovu prislúcha druh slovného druhu, emócia a číselná hodnota. Ďalšou dôležitou informáciou je zistenie počtu slov pre každú emóciu, súčet absolútnych hodnôt číselného ohodnotenia nájdených slov pre emóciu a taktiež zistenie celkového počtu slov nachádzajúcu sa v dokumente.

Druhým krokom je výpočet výsledku, ktorý sa skladá z výpočtu pravdepodobností pre každú jednu emóciu (odpor, hnev, radosť, smútok, strach, úžas). Pravdepodobnosť výskytu jednej emócií vypočítame podľa vzorca:

$$P_i = \frac{n}{m} * 100 * \sum_{j=0}^n t_j$$

kde  $P_i$  je pravdepodobnosť počítanej emócie,  $n$  je počet nájdených slov počítanej emócie,  $m$  je počet všetkých slov v dokumente a  $t_j$  je číselná hodnota slova vyjadrená v absolútnej hodnote.

Posledným krokom je výpočet víťaznej emócií, ktorý pozostáva z výpočtu koeficientu,

$$k = \frac{100}{S_p}$$

kde  $S_p$  je súčet vypočítaných pravdepodobností všetkých emócií. Prepočte všetkých pravdepodobnosti je podľa vzorca:

$$P_j = P_i * k$$

kde  $P_j$  sú výsledné emócie vyjadrené v dokumente a  $k$  je koeficient výpočtu. Výstupom je najvyššia hodnota pravdepodobnosti emócií  $P_j$ , ktorá bola vyjadrená v dokumente. Názorná ukážka hodnoteného slovenského dokumentu je na nasledujúcom obrázku.

## VI. VYHODNOTENIE

Vyhodnotenie efektívnosti klasifikácie sme testovali na 100 článkoch, ktoré boli zamerané na športovú problematiku, všetky články sme získali z internetu zo stránky www.šport.sk. Klasifikácia bola vykonaná na každý jeden článok pomocou slovníkového prístupu. Vyhodnotenie spočíva v stanovení výsledkov presnosti a návratnosti. Pre vyhodnotenie dosiahnutých výsledkov v práci je však potrebné mať vstupné hodnoty pre aplikáciu, jednými s takýchto hodnôt v našej práci sú ohodnotené športové články od aplikácie a od dvoch na sebe nezávislých expertov. Tieto údaje sa potom premietnu do kontingenčnej tabuľky, z ktorej sa vypočíta presnosť a návratnosť.

Tab. 2 Kontigenčná tabuľka.

<b>Druh emócií</b>		<b>Výsledok od Experta</b>	
		<b>Áno</b>	<b>Nie</b>
<b>Výsledok od aplikácie</b>	<b>Áno</b>	<b>A</b>	<b>B</b>
	<b>Nie</b>	<b>C</b>	<b>D</b>

### A. Vyhodnotenie presnosti a návratnosti pri aplikácii a expertovi č.1

Tab. 3 Výsledky presnosti a návratnosti emócií v percentách

Emócia	Odpor	Hnev	Radosť	Smútok	Strach	Úžas
Presnosť	33	60	68	60	50	30
Návratnosť	67	33	64	27	13	40

Pri tomto porovnávaní aplikácie s expertom č.1 sme najlepšiu presnosť hodnotenia dosiahli pri emócií radosť. Pretože najviac článkov hodnotených rovnako aplikáciou a expertom bolo práve pri tejto emócií pomere ku všetkým hodnoteným článkom aplikáciou pri tejto emócií. Najmenšia presnosť sa dosiahla pri emócií úžas, pretože tu bolo najmenej rovnako hodnotených článkov od aplikácie a experta pomere ku všetkým hodnotením článkom aplikáciou vyjadrujúcich emócií úžas.

U návratnosti máme najlepšie výsledky pri emócií odpor, pretože aplikácia sa s expertom pri hodnotení emócie odpor zhodla najviac krát voči všetkým článkom ohodnotených



emóciou odpor. Zase najmenej sa aplikácia zhodla s expertom pri emócií strach zo všetkých možných článkov vyjadrujúcich emóciu strach.

*B. Vyhodnotenie presnosti a návratnosti pri aplikácii a expertovi č.2*

Tab. 4 Výsledky presnosti a návratnosti emócií v percentách

Emócia	Odpor	Hnev	Radosť	Smútok	Strach	Úžas
Presnosť	17	0	51	20	0	24
Návratnosť	50	0	48	60	0	29

Pri tomto porovnávaní aplikácie s expertom č.2 sme najlepšiu presnosť hodnotenia dosiahli znovu pri emócií radosť ako pri expertovi č. 1. Pretože najviac článkov hodnotených rovnako aplikáciou a expertom bolo práve pri tejto emócií pomere ku všetkým hodnoteným článkom aplikáciou pri tejto emócií. Najmenšia presnosť, teda nulová sa dosiahla pri emóciách hnev a strach, pretože aplikácia a expert sa navzájom nezhodli ani pri jednom článku, ktorý by mal vyjadrovať tieto emócie.

U návratnosti máme najlepšie výsledky pri emócií odpor, kde sa aplikácia s expertom pri hodnotení emócie odpor zhodli najviac krát voči všetkým článok ohodnotených emóciou odpor. Zase nulová návratnosť je pri emóciách hnev a strach.

*C. Vyhodnotenie presnosti a návratnosti pri expertovi č.1 a expertovi č.2*

Tab. 6 Výsledky presnosti a návratnosti emócií v percentách

Emócia	Odpor	Hnev	Radosť	Smútok	Strach	Úžas
Presnosť	67	11	64	64	25	50
Návratnosť	100	20	64	39	67	45

Pri porovnaní expertov č.1 a č.2 boli najlepšie výsledky presnosti u emócií odpor a najhoršie pri emócií hnev. Návratnosť bola sto percentná u emócií odpor a najmenšia pri emócií hnev.

*D. Vyhodnotenie presnosti a návratnosti pri aplikácii a expertom*

Tab. 2 Výsledky presnosti a návratnosti emócií v percentách

Emócia	Odpor	Hnev	Radosť	Smútok	Strach	Úžas
Presnosť	25	30	60	40	25	27
Návratnosť	60	21	56	14	9	34

Pri tomto porovnávaní aplikácie s expertom sme najlepšiu presnosť hodnotenia dosiahli pri emócií radosť. Pretože najviac článkov hodnotených rovnako aplikáciou a expertom bolo práve pri tejto emócií pomere ku všetkým hodnoteným článkom aplikáciou pri tejto emócií. Najmenšia presnosť sa dosiahla pri emócií strach a odpor kde bolo najmenej rovnako hodnotených článkov od aplikácie a experta pomere ku všetkým hodnotením článkom aplikáciou vyjadrujúcich tieto emócie.

U návratnosti máme najlepšie výsledky pri emócií odpor, pretože aplikácia sa s expertom pri hodnotení emócie odpor zhodla najviac krát voči všetkým článkom ohodnotených emóciou odpor. Zase najmenej sa aplikácia zhodla s expertom pri emócií strach zo všetkých možných článkov vyjadrujúcich emóciu strach.

## VII. ZÁVER

Jedným z cieľov práce bolo popísať k čomu analýza sentimentu slúži, kde sa najčastejšie využíva. Preto v teoretickej časti bola vysvetlená analýza sentimentu, jej výhody a popísané oblasti jej využitia. Ďalej boli popísané metódy kontrolovaného a nekontrolovaného učenia, ktoré sa pri analýze sentimentu používajú. Základným cieľom tejto práce bolo navrhnúť vhodný slovníkový prístup k analýze emócií v slovenských dokumentoch a otestovať ho na dostatočnej množine dát a vypočítať výsledky presnosti a návratnosti pre všetky emócie. Tento prístup sme testovali na množine dát, ktorá sa skladala zo 100 športových článkov. V týchto článkoch sme zisťovali emócie napr. odpor, hnev, radosť, smútok, strach alebo úžas. Výsledky navrhnutej

aplikácie sme porovnávali s výsledkami na sebe nezávislých dvoch expertov. Po porovnávaní výsledkov sme získali dôležité vstupné údaje pre výpočet presnosti a návratnosti. Dôležitou časťou práce bolo vytvorenie slovníka emocionálnych slov, ktorý sme používali pri slovníkovom prístupe klasifikácie emócií. Na vytvorení tohto slovníka sa podieľalo viacero študentov, ktorým je slovník pomocní taktiež pri ich prácach využívaných analýzu sentimentu. Slovník emocionálnych slov je v rozsahu vyše 19-tisíc slov zoradených v abecednom poradí. Posledným bodom práce bolo analyzovanie vyhodnotených výsledkov presnosti a návratnosti. Tieto výsledky by sa do budúcnosti dali zlepšiť použitím viacerých expertov, zväčšením databázy slovníka o ďalšie emočné slová, alebo použitím iného slovníkového prístupu. Na záver môžeme zhodnotiť že analýza sentimentu je ešte stále vo fáze veľkého rozvoja, a jej metódy sa postupom času stále zlepšujú. Dôležitým nástrojom pri analýze sentimentu do budúcnosti by mohla byť možnosť ručne zmeniť určený sentiment pri jeho špatnom vyhodnotení a tým by sme dosiahli lepšie výsledky analýzy sentimentu.

#### REFERENCIE

- [1] LIU, Bing: Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, May 2012. 167s.
- [2] BURYAN, Jozef: Analýza sentimentu v príspevkoch na sociálnej sieti Twitter: Bakalárska práca. Brno: Masarykova univerzita, Fakulta informatiky, 2013. 36 s.
- [3] MOUTHANI, K. et al.: Sentiment Analysis and Classification Based On Textual Reviews [online]. International Conference on 2013. [cit. 2014-5-9]. Dostupné na internete: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6508366&queryText%3DSentiment+Analysis+and+Classification+Based+On+Textual+Reviews>>.
- [4] NEETHU, M.S.; RAJASREE, R.: Sentiment analysis in twitter using machine learning techniques [online]. India: Department of Computer Science and Engineering College of Engineering. [cit. 2014-5-12]. Dostupné na internete: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6726818&queryText%3DSentiment+Analysis+in+Twitter+using+Machine+Learning+Techniques>>.
- [5] KEKE, Cai et al.: Leveraging Sentiment Analysis for Topic Detection [online]. International Conference on Web Intelligence and Intelligent Agent Technology on 2008.

# Testovanie moderných steganografických metód

<sup>1</sup>Jozef TOMÁŠ, <sup>1</sup>Martin BRODA, <sup>1</sup>Dušan LEVICKÝ

<sup>1</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

jozef.tomas@student.tuke.sk, martin.broda@tuke.sk, dusan.levicky@tuke.sk

**Abstrakt** — Článok obsahuje základné informácie o obrazovej steganografii a zameriava sa na porovnanie steganografických algoritmov v statických obrazoch, ktoré slúžia na utajený prenos informácie v pozadí bežne prebiehajúcej komunikácie. V úvode je vysvetlený pojem steganografický systém a sú zadefinované základné vlastnosti, ktoré by mal takýto systém spĺňať. Ďalšia časť príspevku je venovaná prehľadu súčasných metód obrazovej steganografie a stručnému popisu jednotlivých metód. Hlavným prínosom je testovanie a porovnanie vybraných algoritmov steganografie v obrazoch JPEG z hľadiska veľkosti tajnej správy, ktorú možno do obrazu vložiť, a tiež z hľadiska výpočtovej a časovej náročnosti pri vkladani.

**Kľúčové slová** — steganografia, JPEG, statický obraz, ukrývanie tajnej správy, tajná komunikácia

## I. ÚVOD

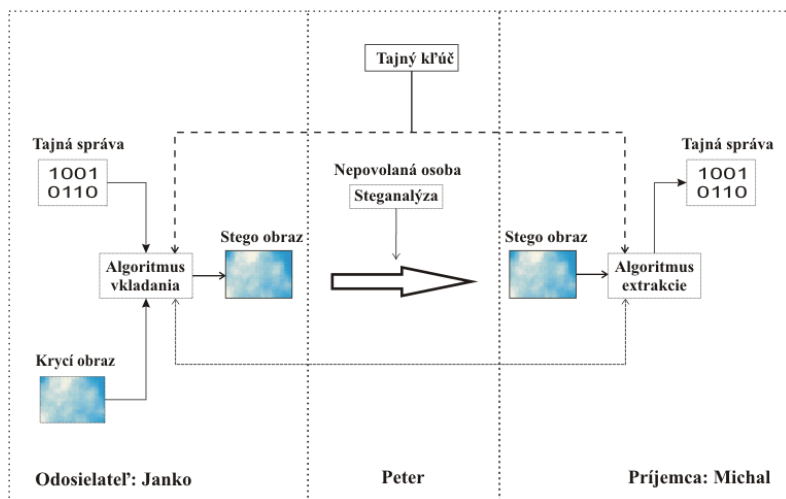
Ochrana informácií a bezpečnosť pri ich prenose informačnými kanálmi v digitálnej podobe je momentálne celosvetovým problémom. Tu sa však priestor pre vývoj vedných disciplín ako je steganografia. Aj keď je steganografia známa už tisícky rokov, pre mnoho ľudí sú pojmy steganografia a steganografické metódy úplne neznámymi pojmami. Slovo steganografia pochádza z gréčtiny a pozostáva z dvoch slov, a to zo slova „steganos“ (skrytý) a „graphein“ (písať). Je to veda o utajovaní komunikácie prostredníctvom ukrytia tajnej informácie v správe [1].

## II. OBRAZOVÁ STEGANOGRAFIA

Steganografia je vednou disciplínou, ktorej úlohou je preniesť tajnú správu v pozadí neutajenej komunikácie. Takáto kamufláž dokáže ochrániť tajnú správu medzi odosielateľom a prijímateľom pred útokmi nepovolaných osôb. Tento prístup sa líši od kryptografie tým, že kladie dôraz na znepriístupnenie komunikácie a nie na utajenie správy. Pre zvýšenie bezpečnosti sa niektoré systémy kryptografie využívajú aj v steganografii.

Zovšeobecný model steganografického systému, ktorý vytvára podprahovú komunikáciu, je možné vidieť na Obr. 1. Hlavným princípom tohto modelu je ukrytie tajnej informácie v digitálnych dátach tak, aby ju tretia strana (nepovolaná osoba) nedokázala odhaliť. Prvým krokom pri vytváraní steganografického modelu je výber krycieho média (cover medium), ktoré je nosičom tajnej správy. Krycie médium môže tvoriť akékoľvek multimédium, napr. text, statické obrazy, audio, video a iné. V druhom kroku sa musí vybrať vhodný algoritmus pre ukrytie dát do krycieho média. Tento algoritmus musí byť rovnaký aj na strane príjemcu. Prostredníctvom algoritmu sa vloží tajná informácia do krycieho média, čím vzniká stego objekt (stego médium). Pre vyššiu bezpečnosť používajú niektoré algoritmy v procese vkladania tajný kľúč.

Na strane príjemcu, ktorý pozná použitý steganografický algoritmus, sa pomocou tajného kľúča, ktorý bol použitý v procese vkladania tajnej informácie, vyextrahuje zo stego objektu tajná správa.



Obr. 1. Principiálna schéma steganografického systému

#### A. Základné požiadavky a rozdelenie steganografických systémov

Na steganografický systém sú kladené nasledovné požiadavky [1]:

- **Nedetekovateľnosť** (Undetectability) - počet vykonaných štatistických zmien, ktoré sú spôsobené vložením tajnej správy.
- **Kapacita** (Capacity) – maximálne množstvo tajnej informácie, ktoré je možné vložiť do krycieho média prostredníctvom danej metódy.
- **Perceptívna transparentnosť** (Perceptual transparency) - schopnosť zachovania si požadovanej kvality daného krycieho média aj po vložení tajnej informácie.
- **Bezpečnosť a robustnosť** (Security and robustness) - odolnosť systému vzhľadom na útoky, znehodnotenie alebo pokusy o odstránenie tajnej správy z označených médií tak, aby nedošlo k zničeniu alebo poškodeniu tajnej správy.

Metódy steganografie v statických obrazoch je možné rozdeliť podľa spôsobu vkladania tajnej informácie do krycieho média. Z tohto hľadiska ich môžeme zatriediť do šiestich základných skupín [2]:

- Substitučné metódy
- Metódy v transformovanej oblasti
- Metódy s rozprestretým spektrom
- Štatistické metódy
- Metódy skreslenia
- Metódy založené na generovaní krycieho obrazu

##### 1) Substitučné metódy

Ich podstata spočíva v zamieňaní nadbytočných častí krycieho obrazu, alebo častí obrazu ktoré sú čo najmenej potrebné, za tajnú informáciu. Zámena je realizovaná tak, aby voľným okom nebolo možné postrehnúť rozdiel na krycom obraze. Aby prijímateľ mohol tajnú informáciu správne dekodovať, musí byť oboznámený o presnej pozícii, kde sa tajná informácia nachádza. Výhodou substitučných metód je ich jednoduchosť. Na druhej strane je ich veľkou nevýhodou bezpečnosť. Tým že sa pri ukrývaní tajnej informácie menia štatistické vlastnosti obrazu je jej odhalenie veľmi jednoduché.

##### 2) Metódy v transformovanej oblasti

Postačujúcu bezpečnosť pre ukrývanie tajných správ nám poskytujú metódy v transformovanej oblasti, pretože ukrývanie informácií vo frekvenčnej oblasti je viac odolné voči narušeniu ako v časovej oblasti. Výhodou týchto metód je ukrývanie tajnej správy do oblastí krycieho obrazu, ktoré sú odolnejšie voči orezávaniu a kompresii. Podmienkou je však malá veľkosť tajnej správy, čím sa zabezpečí čo najmenšia zmena štatistických vlastností obrazu a to je z hľadiska kapacity ich veľkou nevýhodou.

### 3) *Metódy s rozprestretým spektrom*

V komunikačných technológiách sa metódy s rozprestretým spektrom vyvíjali už od roku 1950. Tieto metódy poskytujú veľmi dobrú perceptívnu transparentnosť a taktiež aj veľkú kapacitu pre ukrývanie tajných správ. Ukrytá tajná správa je rozprestretá po celom krycom médiu.

### 4) *Štatistické metódy*

Štatistické metódy sú známe aj ako metódy, ktoré obnovujú štatistické vlastnosti obrazu. Hlavnou úlohou je teda vykonanie čo najmenších zmien štatistických vlastností obrazu. Používanie štatistických metód závisí vo veľkej miere od výberu vhodného krycieho obrazu.

### 5) *Metódy skreslenia*

Metódy skreslenia sa veľmi podobajú substitučným metódam. Rozdiel medzi nimi spočíva v procese extrakcie, kde metódy skreslenia stále potrebujú originálny obraz, do ktorého bola tajná správa vložená. Ich princíp spočíva v doručení stego obrazu prijímateľovi spolu s originálnym obrazom. Na základe ich porovnania prijímateľ extrahuje tajnú správu.

### 6) *Metódy založené na generovaní krycieho obrazu*

Metódy založené na generovaní krycieho obrazu si počas procesu vkladania vytvárajú vlastný krycí obraz. Ich použitie môžeme nájsť v lingvistickej steganografii, kde sa tajná správa ukrýva pomocou textu. Tieto metódy sú veľmi nevýhodné z hľadiska časovej náročnosti, ktorá je pri vkladani a extrakcii tajnej správy veľmi veľká.

## III. PROGRAMOVÁ REALIZÁCIA STEGANOGRAFICKÝCH METÓD A VYTVORENIE DATABÁZY OBRAZOV

Navrhovaný program je určený pre vybrané moderné metódy obrazovej steganografie ako sú nsF5 [3], MB1 [4], MB2 [4], PQ [5], MHF-DZ [6], ktoré pracujú s obrazmi vo formáte JPEG. Pre tento účel bola vytvorená aj databáza obrazov, ktorá bola získaná pomocou rôznych typov fotoaparátov. V navrhnutom programe sa prostredníctvom steganografických metód a databázy krycích obrazov vytvárajú stego obrazy (obrazy s tajnou správou).

### A. *Databáza obrazov*

Na vytvorenie databázy obrazov bolo použitých približne 18000 pôvodných obrazov, ktoré boli nafotené rôznymi typmi fotoaparátov, ako napríklad Nikon D3000, Samsung S730, Sony Ericsson C702, Lumix, Nikon D60, Nikon D3100, Nikon D3200, Olympus FE-115, Olympus X-715. Zo získaných obrazov bola vytvorená množina 2000 obrazov, ktorá bola následne rozdelená do ôsmich skupín, v ktorých boli obrazy tiež zo všetkých fotoaparátov. V každej skupine pozostávajúcej z 250 obrazov som upravil grafické rozlíšenie obrazov na veľkosti 320 x 240 (QVGA), 480 x 320 (HVGA), 640 x 480 (VGA), 800 x 600 (SVGA), 1024 x 768 (XGA), 1600 x 1200 (UXGA) a 1920 x 1080 (HD 1080), aby vzorka pre vkladanie správy bola čo najpestrejšia.

### B. *Programová realizácia*

Ovládanie navrhnutého používateľského prostredia celého programu určeného pre moderné metódy obrazovej steganografie je pre používateľa veľmi jednoduché. Realizované je v programovom prostredí GUIDE v simulačnom programe MATLAB 7.9 R2009b. Program umožňuje výber medzi steganografickými metódami, ktoré sú spomínané v kapitole 3. Po zvolení danej metódy je možné vybrať obrazy JPEG, do ktorých bude tajná správa vložená. Veľkosť tajnej správy je volená na základe parametra *b<sub>pnz</sub>* (*bit per non-zero AC DCT coefficients*), ktorý je v intervale od 0 do 1. Pri testovaní bola do obrazov vkladaná tajná správa o zvolenej veľkosti vo forme vygenerovanej pseudonáhodnej postupnosti bitov.

## IV. EXPERIMENTÁLNE VÝSLEDKY

Prostredníctvom vytvoreného programu bola do krycích obrazov ukrývaná náhodne vygenerovaná tajná správa o veľkosti, ktorá bola ovplyvnená zvoleným parametrom *b<sub>pnz</sub>*. Na základe objektívnych parametrov kvality boli následne porovnávané krycie obrazy s obrazmi, ktoré obsahujú vloženú tajnú správu.

Z objektívnych parametrov kvality bola vybraná stredná kvadratická chyba (MSE), pomocou ktorej bol vypočítaný špičkový pomer signál/šum (PSNR) a maximálnu kapacita

vloženej správy (C). Keďže správa nie je vkladaná do jedného krycieho obrazu ale do vybranej množiny krycích obrazov s rôznou veľkosťou a rozlíšením, hodnoty zobrazené v tabuľkách sú výsledné priemerné hodnoty. V nasledujúcej tabuľke (Tab. 1) sú vypísané priemerné hodnoty vložených bitov tajnej správy do množiny 250 krycích obrazov.

Tab.1 Priemerná veľkosť vlozenej správy (250 obrazov)

<i>bpnz</i>	nsF5, MB1, MB2, MHF-DZ	PQ
	<i>C (kb)</i>	<i>C (kb)</i>
0.05	5,207	4,209
0.25	26,035	21,045
0.5	52,069	42,089
0.75	78,104	63,135
1	104,138	84,179

Ako si môžeme všimnúť, hodnoty pre steganografické metódy nsF5, MB1, MB2 a MHF-DZ sú rovnaké. Hodnoty pri steganografickej metóde PQ sú menšie, pretože veľkosť vkladanej správy sa počíta z obrazov v odtieňoch šedej. Po vypočítaní a následnom vložení tajnej správy boli vypočítané priemerné hodnoty MSE a PSNR, ktoré sú zobrazené v Tab. 2.

Tab.2 Maximálne priemerné hodnoty PSNR , MSE pre vzorku 250 obrazov

<i>bpnz</i>	nsF5		MB1		MB2		PQ		MHF-DZ	
	<i>PSNR (dB)</i>	<i>MSE</i>	<i>PSNR (dB)</i>	<i>MSE</i>	<i>PSNR (dB)</i>	<i>MSE</i>	<i>PSNR (dB)</i>	<i>MSE</i>	<i>PSNR (dB)</i>	<i>MSE</i>
0,02	62,6348	0,05	56,5777	0,19	54,7162	0,28	61,7301	0,05	61,8908	0,05
0,05	57,6474	0,18	52,5344	0,52	50,7751	0,72	56,795	0,19	57,7513	0,14
0,1	54,7038	0,36	50,0629	0,89	48,3463	1,24	50,6386	1,19	54,244	0,31
0,25	49,2409	1,32	45,8122	2,43	44,3585	3,24	45,2348	3,72	49,9873	0,74
0,5	45,1549	3,24	43,0635	4,46	42,0599	5,59	43,7653	4,42	45,8973	1,84
0,75	42,4073	6,43	41,5195	6,65	41,2878	7,04	43,5426	4,67	42,9717	4,01
1	38,2689	15,73	40,5832	7,99	40,5814	7,99	43,0123	4,95	40,4403	8,07

Z tabuľky parametrov vyplýva, že metódy nsF5 a MHF-DZ dosahujú lepšie výsledky pri menšej veľkosti vlozenej tajnej správy. Z hľadiska perceptívnej nevnímateľnosti vyhovujú všetky metódy, keďže rozdiel medzi krycím a stego obrazom pri maximálnej veľkosti vlozenej tajnej správy je približne 40 dB. Takýto rozdiel medzi obrazmi je ľudským okom nepostrehnuteľný.

Následne boli steganografické metódy porovnávané z časového hľadiska. Teda ako dlho trvalo steganografickým metódam vkladanie tajnej správy do množiny 250 krycích obrazov. Pri tomto meraní vplýva na kvalitu výsledkov viacero faktorov, ako sú napríklad používaný operačný systém, výkon procesora, počet spustených aplikácií, procesy bežiacie v pozadí a ďalšie. Aby boli výsledky objektívnejšie, meranie bolo vykonané 3-krát a do Tab. 3 bol uvedený najkratší čas vkladania správy.

Tab.3 Porovnanie metód z hľadiska času pri procese vkladania tajnej správy

<i>bpnz</i>	nsF5	MB1	MB2	PQ
0.02	25.19	19:27.37	55:29.84	12:35.39
0.05	26.05	16:16.31	53:52.71	10:53.34
0.1	27.99	16:45.43	1:01:52.18	13:54.31
0.25	28.31	20:29.90	1:03:45.82	12:49.47
0.5	27.87	22:13.56	1:54:32.57	13:06.11
0.75	37.53	23:53.94	1:55:57.12	12:21,12
1	36.64	22:27.45	1:56:47.75	13:10.33

## V. ZÁVER

V dnešnej dobe je čoraz náročnejšie zabezpečiť svoje informácie pri digitálnom prenose. Práve oblasť steganografie poskytuje riešenie pre tento problém. Za posledných pár rokov bol vyvinutý a zdokonalený veľký počet steganografických metód a preto je pri voľbe vhodnej

steganografickej metódy nevyhnutné poznať ich charakteristické vlastnosti. Tento článok je venovaný opisu steganografie v statických obrazoch, moderným steganografickým metódam, ktoré vkladajú tajnú správu do obrazov vo formáte JPEG. Hlavným cieľom bolo porovnať moderné steganografické metódy z hľadiska kapacity pre vloženie tajnej správy a z hľadiska perceptívnej nevnímateľnosti.

#### REFERENCES

- [1] LEVICKÝ, Dušan: Kryptografia v informačnej a sieťovej bezpečnosti. 1. Vyd. Košice: Elfa, s.r.o., 286 s. ISBN 978-80-8086-163-6.
- [2] NAGHAM, Hamid – ABID, Yahya – BADLISHAH, Ahmad – OSAMAH, M.: Image Steganography Techniques: An Overview. Internal Journal of Computer Science and Security, Volume-6, Issue-3, 2012.
- [3] FRIDRICH, J. – GOLJAN, M. – SOUKAL, D.: Wet paper codes with improved embedding efficiency, [online], Information Forensics and Security, IEEE Transactions on, vol.1, no.1, pp.102,110, March 2006
- [4] FRIDRICH, J. – GOLJAN, M. – SOUKAL, D.: Perturbed quantization steganography. ACM Multimedia System Journal, 11(2): pp.98-107, 2005.
- [5] SALLEE, P.: "Model-based methods for steganography and steganalysis," International Journal of Image and Graphics, 167-190 (2005).
- [6] BANOCI, V. – BUGAR, G. – LEVICKÝ, D. – KLENOVICOVA, Z.: Histogram secure steganography system in JPEG file based on modulus function, [online], Radioelektronika (RADIOELEKTRONIKA), 22nd International Conference, pp.1-4, 17-18 April 2012. ISBN: 978-1-4673-0659-1.

# SDI-12 adaptér pre monitorovanie dát zo senzorov kvality vody

<sup>1</sup>Marek LABAN, <sup>2</sup>Miloš DRUTAROVSKÝ

<sup>1,2</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky Technickej univerzity v Košiciach

<sup>1</sup>marek.laban@student.tuke.sk, <sup>2</sup>milos.drutarovsky@tuke.sk

**Abstrakt** – Článok opisuje adaptér umožňujúci monitorovať a zaznamenávať dáta zo senzorov s integrovaným rozhraním SDI-12. Adaptér tvorí základ vyvíjaného zariadenia pre pripojenie inteligentných senzorov na monitorovanie kvality vody. Jadro adaptéra tvorí minimodul, na ktorom je osadený 8-bitový mikroprocesor C8051F381 firmy Silicon Laboratories s integrovaným USB rozhraním. Procesor zabezpečuje prenos dát medzi dvoma sériovými rozhraniami s rozdielnymi prenosovými rýchlosťami, ako aj záznam monitorovaných dát do lokálnej 128-Mbitovej FLASH pamäte pripojenej cez SPI rozhranie. Príspevok opisuje vytvorený hardvér, testovací softvér a celkovú funkcionálnosť vytvoreného zariadenia.

**Kľúčové slová** – adaptér, SDI-12, Silabs, mikrokontrolér, senzor, kvalita vody, UART, USB

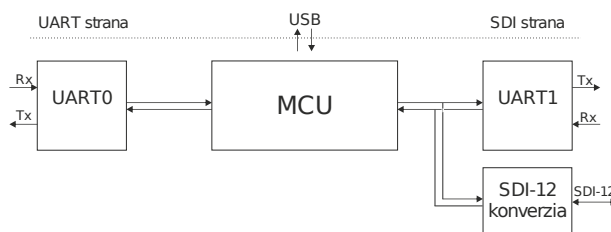
## I. ÚVOD

Realizácia opisovaného SDI adaptéra vyplynula z požiadavky vývoja monitorovacieho zariadenia senzorov kvality vôd, ktoré je budované na Katedre elektroniky a multimediálnych komunikácií FEI TU Košice v rámci medzinárodného projektu HUSK. Inteligentné senzory pre monitorovanie kvality vody typicky využívajú štandardizované komunikačné rozhranie SDI-12 [1]. Cieľom bolo vyvinúť malý modul, ktorý umožní realizovať autonómny záznam dát poskytovaných inteligentnými senzormi, realizovať predspracovanie dát a zabezpečiť pripojenie nadradených systémov prostredníctvom štandardizovaných rozhraní UART a USB.

## II. HARDVÉROVÁ ČASŤ SDI ADAPTÉRA

### A. Požiadavky na SDI adaptér

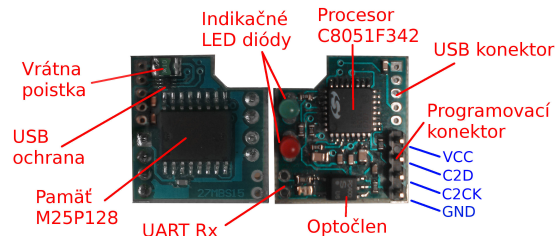
Základnou funkciou adaptéra je realizácia prevodníka UART - SDI-12 (obr. 1), pričom rozhranie SDI-12 vykonáva prenosy s rýchlosťou 1200 baudov (rýchlosť je pre SDI rozhranie štandardizovaná) a rozhranie UART prenosy s vyššou rýchlosťou, napríklad 115200 baudov. Vytvorený softvér preto musí realizovať komunikáciu medzi dvoma sériovými linkami s rôznymi prenosovými rýchlosťami za pomoci softvérového riadenia toku (XON, XOFF). Prevodník ma tiež pre potreby vývoja umožňovať alternatívnu komunikáciu prostredníctvom rozhraní UART - UART. Celá komunikácia ma byť súčasne zaznamenávaná do dostatočne veľkej FLASH pamäte, z ktorej sa pomocou USB rozhrania následne môžu zaznamenané dáta vyčítať do nadradeného systému.



Obr. 1: Bloková schéma prenosu dát prostredníctvom SDI-12 adaptéra a jeho rozhraní

Hardvérovým jadrom SDI adaptéra sa pôvodne mal stať minimodul (obr. 2) s osadeným 8-bitovým jednočipovým procesorom C8051F342 firmy Silicon Laboratories (Silabs). Tento minimodul bol v minulosti na KEMT vyvinutý pre inú aplikáciu, avšak obsahoval FLASH pamäť M25P128 [2] s veľkosťou 128 Mbitov, ktorú bolo možné využiť pre zaznamenávanie komunikácie.





Obr. 2: Pôvodný minimodul s procesorom C8051F342 firmy Silabs - vľavo spodná strana, vpravo vrchná strana

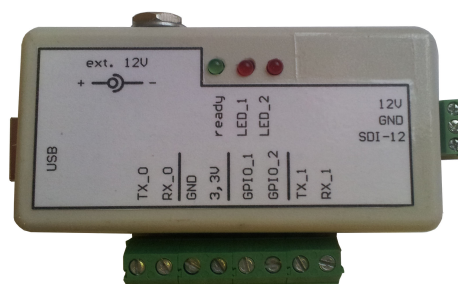
### B. Hardvérové riešenie SDI adaptéra

Vyvinutý hardvér pozostáva z minimodulu a vytvorenej základnej dosky plošných spojov (DPS), ktorá rozširuje možnosti zapojenia jednotlivých sériových komunikačných liniek. Osadený minimodul obsahuje oproti jeho pôvodnej verzii niekoľko zmien. Pôvodný procesor typu C8051F342 bol z dôvodu absencie jedného sériového rozhrania nahradený procesorom C8051F381 [3], ktorý je programovo a vývodovo plne kompatibilný. Ďalšou podstatnou zmenou je spôsob zapojenia FLASH pamäte M25P128, na ktorej bol deaktivovaný vývod pre pozastavenie prenosu (Hold) a vývod používaný na ochranu pred zmenou v chránenej oblasti pamäte (Write protect).

Navrhnutý SDI adaptér môže pracovať v režime UART0 - UART1 alebo UART0 - SDI-12. Umožňuje pripojiť 2 sériové linky (obr. 3), pričom každá sériová linka využíva vývody Tx a Rx. Ďalej má adaptér vyvedené 2 všeobecné vstupno-výstupné vývody (GPIO), ktorými je možné v procesore vyvolať prerušenie. Tieto GPIO vývody sú určené na hardvérové riadenie toku dát pre UART0 pokiaľ by bolo v aplikácii požadované.

Na DPS sú osadené súčiastky potrebné pre transformovanie UART rozhrania na SDI-12 rozhranie. Keďže rozhranie SDI-12 používa iba jeden vodič na obojsmernú sériovú komunikáciu, nastavenie príjmu a vysielača sa uskutočňuje prostredníctvom jedného GPIO vývodu. Pre napájanie senzora prostredníctvom rozhrania SDI-12 (12 V, dáta, zem) sú vyvedené všetky potrebné vývody.

Z napätového regulátora procesora je pre všeobecné účely vyvedený aj vývod s úrovňou 3.3 V. SDI adaptér je možné napájať z elektrickej siete pomocou externého napájacieho adaptéra so stabilizovaným napájacím napätím 12 V, alebo prostredníctvom USB rozhrania (vývod na 12 V je vtedy neaktívny). Výber napájania ako aj výber rozhrania SDI-12/UART1 je možné vykonať prostredníctvom nastavení konfiguračných vývodov vo vnútri adaptéra (obr. 4). Celkovo sú vyrobené a otestované 3 kusy SDI adaptéra, v každom z nich je nahratý zavádzací program pre aktualizáciu aplikačného firmvéru prostredníctvom USB rozhrania.

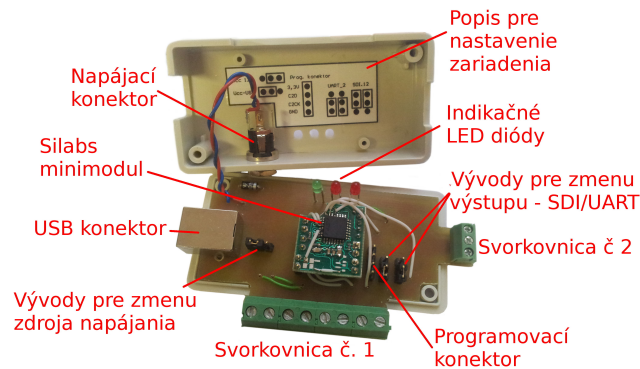


Obr. 3: SDI-12 adaptér - vonkajší pohľad

### C. Procesor C8051F381

Procesor C8051F381 spoločnosti Silicon Laboratories [3] je 8-bitový mikroprocesor s integrovaným USB rozhraním založený na jadre procesora Intel 8051. Na rozdiel od procesora Intel 8051 však využíva architektúru prekrývania strojových inštrukcií (pipelined instruction architecture) s ktorou je v 1-2 taktach procesora vykonaných 70% inštrukcií.

Tento procesor patrí do rodiny C8051F38x procesorov, ktorá predstavuje novšiu, vylepšenú, plnú náhradu za rodinu C8051F34x s ktorou je vývodovo a programovo kompatibilná. Medzi inovácie tejto rodiny patrí integrovaný druhý UART do všetkých MCU, ďalej je to druhá SMBUS zbernica a taktiež ďalšie dva 16-bitové časovače. Zaujímavá je však ich cena, pretože v porovnaní so staršou rodinou je približne polovičná.



Obr. 4: SDI-12 adaptér - vnútorné usporiadanie

Firma Silicon Labs poskytuje bohaté množstvo aplikačných príručiek, knižníc ako aj príkladov potrebných jednoduchú implementáciu USB rozhrania. Medzi osvedčené riešenia patrí najmä vývojová sada USBXpress, ktorá poskytuje všetky softvérové prostriedky potrebné pre sprevádzkovanie USB rozhrania.

Procesor má vstavaný interný 48 MHz oscilátor, vďaka ktorému hardverová implementácia USB rozhrania nepotrebuje napojenie ďalšieho externého oscilátora. Ďalej poskytuje viaceré komunikačné rozhrania ako sú I2C, SPI, UART, SMBUS. Medzi výhody patrí prepínacie pole (crossbar), ktoré do istej miery umožňuje priradiť danému rozhraniu užívateľom zvolený konkrétny vývod. Ako nevýhoda sa môže javiť absencia DA prevodníka a taktiež relatívne malá interná pamäť RAM.

Základné vlastnosti procesora C8051F38x:

- interný 48 MHz a 80 kHz oscilátor,
- programová pamäť 64 kB (FLASH), dátová pamäť RAM 4352 B,
- 1kB USB FIFO register (môže byť použitý ako dátová pamäť RAM),
- SPI, 2x SMBUS/I2C, 2x UART, USB rozhranie s 1kB FIFO registrom,
- napätový regulátor, napätová referencia,
- AD prevodník, senzor teploty, analógový komparátor,
- 6 x 16-bitový časovač,
- vstupno výstupné vývody na všeobecné použitie: 40 (obsahuje aj EMIF)/25,
- dostupné púzdra: LQFP32 (9x9 mm), QFN32 (5x5 mm), TQFP48 (9x9 mm).

### III. SOFTVÉROVÁ ČASŤ SDI ADAPTÉRA

#### A. Firmvér SDI adaptéra

Vytvorený firmvér v prvom rade zabezpečuje bezstratový prenos dvoch sériových asynchronných rozhraní (UART-ov). Pre tento účel sú dáta na príjímачích stranách adaptéra (obr. 1) ukladané do 256-bitového zásobníka. LED diódy umiestnené na adaptéri indikujú, že v tomto zásobníku sa nachádzajú neodoslané dáta. Zásobník je chránený pred pretečením pomocou softvérového riadenia toku na strane UARTu. Na strane SDI takéto riadenie nie je, pretože sa predpokladá pomalší prenos ako na strane UARTu. Časť softvéru potrebnú pre komunikáciu prostredníctvom SDI-12 bude súčasťou aplikačných programov a je ju možné do adaptéra nahrávať pomocou zavádzacieho programu (bootloadera) cez USB rozhranie.

Celá komunikácia je zaznamenávaná do FLASH pamäte M25P128. Využíva sa pri tom knižnica vytvorená a opísaná v [4].

Rozhranie USB rozširuje celkové možnosti zariadenia, preto na tieto účely bola vytvorená PC aplikácia. USB konektivita na strane mikroprocesora, ako aj na strane PC je zabezpečená prostredníctvom vývojovej sady USBXpress. Na strane mikroprocesora je použitá knižnica USBXpress, ktorá umožňuje prácu s rozhraním bez hlbšieho štúdia katalógového listu procesora.

Firmvér môže byť do mikroprocesora naprogramovaný prostredníctvom špeciálneho programovacieho adaptéra, alebo prostredníctvom USB rozhrania za pomoci zavádzacieho programu. Zavádzací program (bootloader) zvyčajne označuje časť firmvéru, ktorá sa nachádza v programovej pamäti procesora a má schopnosť prijímať nové aktualizácie prostredníctvom dostupného rozhrania. V tomto prípade je to rozhranie USB, čím je v ďalšom vývoji umožnená bezproblémová aktualizácia firmvéru bez nárokov na špecializovaný hardvér či skúsenosti používateľa. V rámci tohoto projektu boli vytvorené 2 verzie programu, prvá pre špecializovaný adaptér a druhá pre zavádzací program.

Dokumentácia programu je v rozsahu postačujúcom na bezproblémové zvládnutie a pochopenie všetkých algoritmov.

Funkcionalita firmvéru:

- zabezpečuje komunikáciu medzi dvoma sériovými linkami,
- na strane UART0 sa používa softvérové riadenie toku,
- zaznamenáva komunikáciu do FLASH pamäte,
- zabezpečuje komunikáciu prostredníctvom USB rozhrania pomocou knižnice USBXpress,
- obsahuje sadu príkazov, ktoré je možné zadávať prostredníctvom PC aplikácie.

#### B. PC aplikácia SDI adaptéra

PC aplikácia je vytvorená v jazyku C. Predstavuje užívateľské rozhranie vytvorené pre komunikáciu s SDI adaptérom prostredníctvom rozhrania USB. Na zabezpečenie USB komunikácie bola použitá knižnica USBXpress *SiUSBXp.lib* [5], ktorej hlavičkový súbor bol čiastočne upravený (úpravy sú zaznamenané v samotnom súbore). Napriek tomu, že knižnica je stavaná pre jazyk C++, neboli zaznamenané problémy s nekompatibilitou.

Vytvorené prostredie sa javí ako jednoduchý terminálový program (príkazový riadok). Užívateľ v ňom zadáva jednoduché príkazy, napríklad pre prezeranie zaznamenaných dát z FLASH pamäte SDI adaptéra, zmenu prenosovej rýchlosti sériového rozhrania a podobne. Tieto príkazy je možné ďalej dopĺňať a vytvoriť tak ďalšie operácie pre prácu s cieľovým zariadením.

Pre vývoj tejto aplikácie bolo použité prostredie Dev-C++. Je zamerané na programovanie v jazyku C/C++, jeho výhodou je malá veľkosť inštalovanej distribúcie na disku a dostupnosť zdarma. Prednostne používa voľne prístupný (GNU) gcc kompilátor [6]. Použitá verzia mala označenie Orwell Dev-C++ v5.3 [7].

Funkcionalita PC aplikácie:

- umožňuje zmenu prenosovej rýchlosti sériovej linky (UART0),
- umožňuje uloženie histórie prenášaných dát SDI adaptéra,
- umožňuje zmazanie celej FLASH pamäte,
- informuje o stave zaplnenosti pamäte,
- umožňuje aktiváciu zavádzacieho programu.

### IV. TESTOVANIE SDI ADAPTÉRA

#### A. Test spoľahlivosti prenosu SDI adaptéra

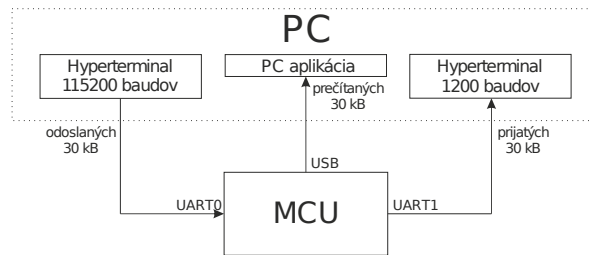
Funkčnosť celého zariadenia SDI adaptéra spočíva v zabezpečení komunikácie medzi dvoma sériovými rozhraniami s rôznymi prenosovými rýchlosťami, riadená pomocou softvérového riadenia toku, ktorá má byť súčasne zaznamenaná do pamäte FLASH. Najjednoduchší spôsob ako to overiť spočíval v napojení dvoch konverzných FTDI káblov [8] na obe sériové rozhrania SDI adaptéra. Ďalej bolo potrebné v programe Hyperterminal [9] nastaviť správne prenosové rýchlosti pre jednotlivé FTDI prevodníky (SDI - 1200 baudov a UART - 115200 baudov) a vzájomne preposielať znaky. Následne pomocou vytvorenej PC aplikácie prečítať záznam prebiehajúcej komunikácie.

Pre zvýšenie zaťaženia SDI adaptéra bolo potrebné vysielat' intenzívny tok dát (Obr. 5). Preto sa v hypertermináli pre jeden FTDI kábel nastavilo zachytávanie textu pomocou menu *Prenos - Zachytávať text*, a pre druhý FTDI kábel odosielanie textového súboru pomocou menu *Prenos - Odoslať textový súbor*. Textový súbor bol vytvorený v poznámkovom bloku mnohonásobným skopírovaním jednej frázy (napr. `"/** Hello World!!` "). Takýmto spôsobom je možné vytvoriť súbor s veľkosťou niekoľko megabajtov, na testovanie však postačil súbor s veľkosťou niekoľko desiatok kilobajtov.

Test dopadol podľa očakávaní, odoslaný súbor bol prijatý v rovnakej veľkosti s rovnakým obsahom na druhej strane, záznam prebiehajúcej komunikácie bolo taktiež možné vyčítať v plnom obsahu prostredníctvom USB rozhrania. Odosielanie textového súboru bolo vykonané pre obe strany sériového rozhrania. Softvérové riadenie toku fungovalo korektne. Pri vývoji aplikačného firmvéru je potrebné zohľadniť to, že do FLASH pamäte sa zapisuje blokovo každých 128 bajtov, preto je potrebné preposlať ešte ďalšie znaky do dĺžky 128 bajtov, prípadne vhodné vyplniť chýbajúce dáta.

#### B. Testovanie celkovej rýchlosti zariadenia SDI adaptér

V závere bolo vykonané meranie na zrealizovanom reálnom zariadení. Dáta boli vyčítavané z pamäte a prenášané priamo do hostovského počítača, výsledky sú uvedené v tabuľke 1. Vyplýva z nich, že zvýšením taktovacej frekvencie procesora sa môže skrátiť čas prenosu zariadenia na polovicu. Tento rozdiel je spôsobený práve SPI rozhraním externej FLASH pamäte. Meranie bolo zopakované aj pri použití zavádzacieho programu, výsledná prenosová rýchlosť ostala nezmenená.



Obr. 5: Princíp vykonaného testu na príklade prenosu súboru s veľkosťou 30 kB.

Počas testovania došlo k zaujímavému javu, keď pri napájaní notebooku (ASUS K53E - Win 7 Ultimate) z batérie došlo počas merania rýchlosti prenosu k poklesu rýchlosti z 236 KB/s na 150 KB/s. Tento jav nebol podrobený bližšiemu skúmaniu, pri programovaní cieľovej aplikácie je preto potrebné otestovať funkcionálnosť nadradeného zariadenia aj pri batériovom napájaní.

Tabuľka 1: Prehľad dosiahnutých prenosových rýchlostí pri prenose súboru z FLASH pamäte do hosťovského počítača

Veľkosť súboru	1842 KB	
Frekvencia procesora	24 MHz	48 MHz
Čas stiahnutia	15,3 s	7,8 s
Dosiahnutá rýchlosť	120 KB/s	236 KB/s

## V. ZÁVER

Navrhnutý a zrealizovaný SDI-12 adaptér umožňuje vývoj a testovanie aplikačných algoritmov pre ďalšie spracovanie údajov získaných z inteligentných senzorov. Vytvorený softvér sprostredkúva a zaznamenáva komunikáciu medzi dvoma sériovými rozhraniami s rôznymi rýchlosťami. Vyvinutá terminálová PC aplikácia sprostredkúva komunikáciu medzi mikrokontrolérom a PC prostredníctvom moderného USB rozhrania.

Všetky funkčné časti SDI adaptéra boli otestované a zistené problémy boli odstránené. Funkčnosť SDI adaptéra však ešte nebola testovaná v prevádzkových podmienkach, keďže v čase jeho finalizácie neboli ešte dostupné objednané inteligentné multisondy.

Hlavným prínosom tohoto projektu je vytvorený modul, ktorý umožňuje ďalší vývoj algoritmov finálneho monitorovacieho zariadenia, ako aj doposiaľ vyvinuté algoritmy na prenos prostredníctvom sériových rozhraní a USB rozhrania mikrokontroléra. PC aplikácia vytvorená pre USB rozhranie zariadenia môže byť po istých obmenách použitá aj v rámci iných projektov, využívajúcich knižnicu USBXpress pre 8-bitové procesory spoločnosti Silicon Laboratories.

## POĎAKOVANIE

Tento článok bol vytvorený v rámci realizácie projektu Wireless Sensor Network for wATER QuALity Monitoring - WSN-AQUA (reg. číslo: HUSK/1101/1.2.1/0091) a realizácie projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] "Sdi-12 a serial-digital interface standard for microprocessor-based sensors," 2013. [Online]. Available: <http://www.sdi-12.org/current%20specification/SDI-12`version1`3%20January%2026,%202013.pdf>
- [2] "M25P128 Datasheet." [Online]. Available: <http://www.chipfind.net/datasheet/pdf/stmicroelectronics/m25p128.pdf>
- [3] "C8051F38x MCU." [Online]. Available: <http://www.silabs.com/products/mcu/usb/pages/c8051f38x.aspx>
- [4] M. Laban, "Jednočipové aplikácie s usb konektivitou," bakalárska práca, KEMT FET TU Košice, jún 2014.
- [5] "Usbexpress dev. tools." [Online]. Available: <http://www.silabs.com/products/mcu/Pages/USBXpress.aspx>
- [6] "GNU Compiler Collection." [Online]. Available: <https://gcc.gnu.org/>
- [7] "Orwell dev-c++ site." [Online]. Available: <http://orwelldevcpp.blogspot.sk/>
- [8] "Ttl-232r datasheet." [Online]. Available: <http://www.ftdichip.com/Support/Documents/DataSheets/Cables/DS`TTL-232R`CABLES.pdf>
- [9] "Hyperterminal for windows 7." [Online]. Available: <http://www.kmtronic.com/hyperterminal-for-windows-7.html>

# Predspracovanie dát z UWB radaru pomocou Mini PC MK802 III

<sup>1</sup>Jaroslav HUTIRA, <sup>2</sup>Miloš DRUTAROVSKÝ

<sup>1,2</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky Technickej univerzity v Košiciach

<sup>1</sup>jaroslav.hutira@student.tuke.sk, <sup>2</sup>milos.drutarovsky@tuke.sk

**Abstrakt** – Cieľom príspevku je ukázať možnosti využitia multimediálneho počítača Mini PC MK802 III ako výpočtového jadra na predspracovanie dát v senzorovej sieti UWB radarov. V článku je opísaný použitý modul s výkonným dvoj-jadrovým procesorom ARM Cortex-A9, použitá technika predspracovania dát pre UWB radar využívajúci 4095 prvkovú M-postupnosť, dosiahnuté výsledky a čas trvania predspracovania dát, čo predstavuje jednu z kritických operácií, ktoré je potrebné realizovať vo vložnom senzorovom uzle na báze UWB radaru.

**Kľúčové slová** – ARM Cortex-A9, FFT, Mini PC, MK802 III, M-postupnosť, UWB radar, senzorová sieť

## I. ÚVOD

Mini počítač MK802 III [1] s výkonným dvoj-jadrovým procesorom ARM Coterx-A9 [2] je primárne určený na multimediálne účely, avšak ponúka dostatočne veľký výpočtový výkon, ktorý je možné s použitím vhodného Operačného Systému (OS) a jeho optimalizáciou využiť aj na spracovanie dát z UWB radaru. To však vyžaduje optimalizáciu použitého programu a výber vhodných algoritmov predspracovania. Predspracovanie dát z UWB radaru, ktorý využíva M-postupnosť sa typicky realizuje vo frekvenčnej oblasti s využitím rýchlej Fourierovej transformácie. Využitie signálov ktorých dĺžka nie je mocninou 2 vyžaduje špecifický výber vhodnej FFT knižnice.

## II. MINI POČÍTAČ MK802 III

Mini PC MK802 III Obr. 1 je produktom spoločnosti Rikomagic [1] a je tretou verziou tejto rady Mini PC. Toto zariadenie predstavuje moderné a výkonné riešenie v oblasti multimediálnych systémov. Jedná sa o počítač, nie o veľa väčší ako kreditná karta ( $40\text{ mm} \times 97\text{ mm} \times 8\text{ mm}$ ), ktorý obsahuje dvoj-jadrový procesor RK3066 s jadrami ARM Cortex-A9 taktovanými frekvenciou 1,6 GHz a pamäťou RAM 1 GB typu DDR3. Uvedené parametre ponúkajú vysoký výpočtový výkon a spolu so správnym softvérovým vybavením je možné Mini PC využiť aj na splnenie komplexnejších úloh [3].



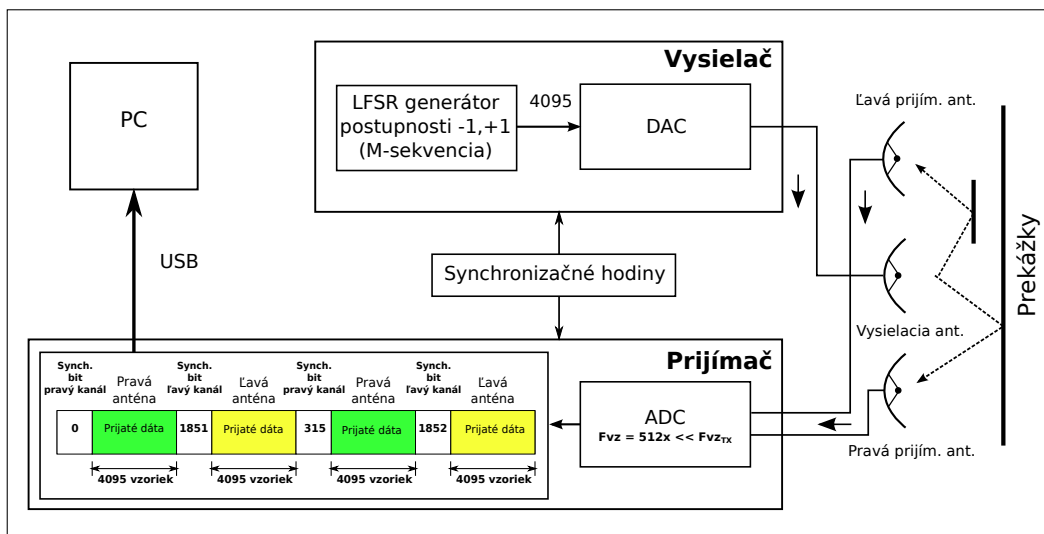
Obr. 1 Mini počítač MK802 III s dvoj-jadrovým procesorom ARM Cortex-A9.

Na Obr. 1 je znázornený Mini PC MK802 III, ktorý výrobca distribuoval s OS Android vo verzii 4.1 Jelly Bean. Vzhľadom na to, že OS Android je určený na multimediálne účely, bol nahradený OS Linux Picuntu [4], ktorý vychádza zo známej distribúcie Ubuntu 12.10 Quantal armhf.

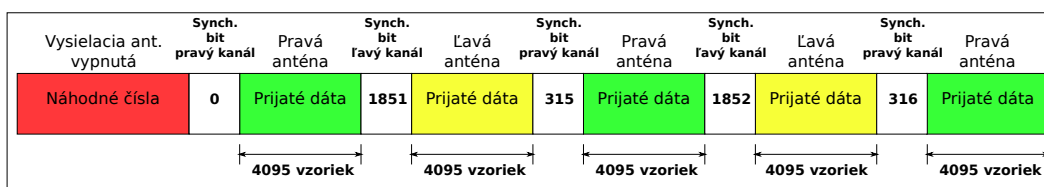


### III. SPRACOVANIE SIGNÁLU V UWB RADARE NA BÁZE M-POSTUPNOSTI

Na Obr.2 je znázornená bloková schéma UWB radaru na báze M-postupnosti. Radar tvorí jeden prijímač a dva vysielajúce, pričom radar využíva jednu vysielaciu a dve prijímacie antény. Dáta sa po prijatí a základnom spracovaní v technických prostriedkoch radaru (typicky na báze obvodov FPGA) posielajú cez USB rozhranie do výpočtového uzla (zvyčajne počítača) kde sa spracujú.



Obr. 2 Princíp generovania, vysielania a prijímania signálu v UWB radare s M-postupnosťou.



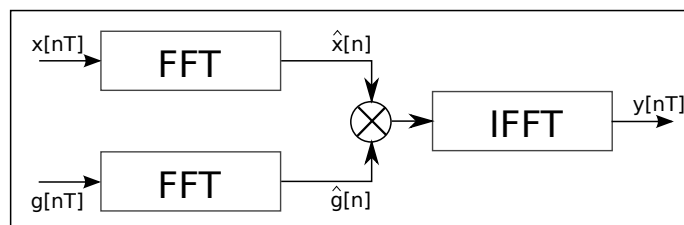
Obr. 3 Prijaté a digitalizované dáta z prijímacích antén usporiadané do paketov.

Vo vysieláči UWB radaru, ktorý sme mali k dispozícii, sa generuje LFSR (angl. Linear Feedback Shift Register) generátorom 4095-prvková periodická M-postupnosť (angl. M-sequence) prvkov  $\pm 1$ . Vygenerované vzorky sa cez číslicovo-analógový prevodník (DAC) posielajú na vysielaciu anténu, táto vyšle UWB signál do priestoru, kde sa od prekážok odrazí a s rôznym oneskorením sa vráti do prijímacích antén. Z prijímacích antén sa signál posiela do prijímača, kde sa digitalizuje analógovo-číslivým prevodníkom (ADC), ktorého vzorkovacia frekvencia  $F_{VZ}$  je 512-krát menšia, ako vzorkovacia frekvencia DAC vo vysieláči, čím vzniká prevzorkovanie signálu. Takto digitalizovaný signál je usporiadaný do postupnosti paketov ako je znázornené na Obr. 3, kde sa striedavo usporiadávajú postupnosti so 4095 vzorkami z oboch prijímacích antén. Každá 4096-ta vzorka je synchronizačná a určuje ktorej anténe prislúchajú dáta. Takto usporiadané dáta sa posielajú prostredníctvom USB rozhrania do počítača na spracovanie. Celý proces sa opakuje každých približne 70 ms.

### IV. ALGORITMUS PREDSPRACOVANIA DÁT

Vzhľadom na to, že predspracovanie prijatých dát sa realizuje vo frekvenčnej oblasti, používa sa rýchla Fourierova transformácia (FFT). Aktuálne sa na katedre elektroniky a multimediálnych telekomunikácií používajú na predspracovanie a spracovanie programy Matlab a LabVIEW, ktoré využívajú FFT. Rovnaké algoritmy sme implementovali aj do Mini PC. Princíp predspracovania dát je zobrazený na Obr. 4. Programy pre Mini PC MK802 III a OS Linux [3] boli napísané v jazyku C a na rýchlu Fourierovu transformáciu boli zvolené dve vysoko optimalizované knižnice MIXFFT [5] a FFTW [6].

Prijatá postupnosť dát z UWB radaru  $x[nT]$  sa transformuje použitím priamej FFT do frekvenčnej oblasti čím získame vzorky  $\hat{x}[n]$ . Pri predspracovaní dát sa používajú aj referenčné dáta, v tomto



Obr. 4 Princíp predspracovania prijatých dát z UWB radaru vo frekvenčnej oblasti použitím FFT.

prípade ich tvorí ideálna M-postupnosť, ktorú v programe generujeme softvérovým LFSR generátorom. Táto  $g[nT]$  postupnosť tvorená  $\pm 1$  sa transformuje priamou FFT do frekvenčnej oblasti čím získame  $\hat{g}[n]$ . Vo frekvenčnej oblasti na vzorky  $\hat{x}[n]$  a  $\hat{g}[n]$  vynásobia ako komplexné čísla a výsledok sa transformuje pomocou spätnej rýchlej Fourierovej transformácie do časovej oblasti. Vzhľadom na to, že jeden blok dát (kompletná 4095-prvková impulzová odpoveď) prichádza z UWB radaru každých približne 70 ms je nevyhnutné, aby predspracovanie bolo časovo veľmi nenáročné.

#### V. PARALELNÉ PREDSPRACOVANIE PRIJATÝCH DÁT Z UWB RADARU

Definícia diskretnej Fourierovej transformácie (DFT) a rôzne FFT algoritmy používajú komplexné čísla a komplexnú aritmetiku. Z toho vyplýva, že pre transformáciu reálnej postupnosti by mala postačovať len polovičná veľkosť transformácie dát. Neexistuje žiadny spôsob, ako sa zbaviť potreby použitia komplexnej aritmetiky, ale existuje niekoľko spôsobov, ako získať lepší výkon z FFT algoritmu aplikovaného na reálnu postupnosť [7]. Dôležitou vlastnosťou transformácie reálnej postupnosti je symetria jej spektrálnych zložiek vyjadrenú vzťahom.

$$F(k) = F(N - k)^* \quad (1)$$

Túto vlastnosť môžeme využiť na zrýchlenie výpočtu tak, že jednu N-bodovú komplexnú FFT môžeme využiť pre výpočet dvoch N-bodových reálnych FFT. Majme reálne postupnosti  $x(n)$  a  $y(n)$  a definujeme komplexnú postupnosť  $z(n)$ :

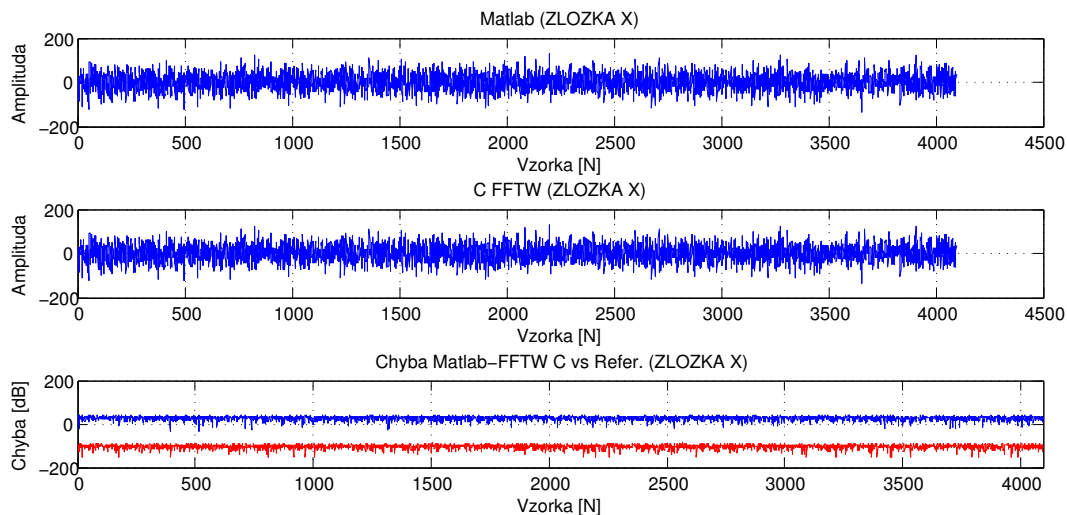
$$z(n) = x(n) + y(n) * j \quad (2)$$

Po rýchlej Fourierovej transformácii postupnosti  $z(n)$  možno extrahovať postupnosti  $x(n)$  a  $y(n)$ . Tento spôsob zlúčenia dvoch reálnych postupností dát prijatých z UWB radaru do jednej komplexnej postupnosti nám umožňuje predspracovať dvojnásobné množstvo dát.

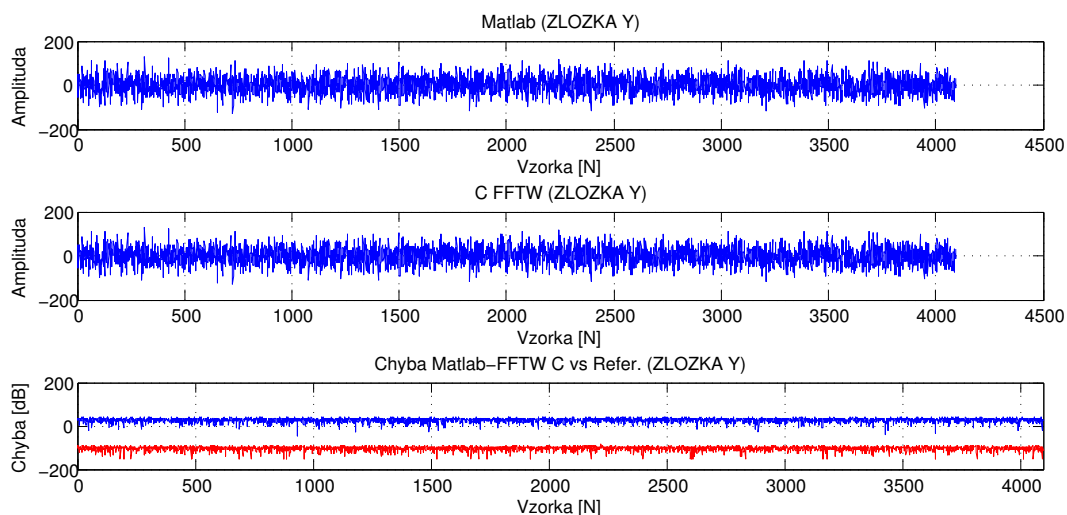
#### VI. EXPERIMENTÁLNE VÝSLEDKY PREDSPRACOVANIA DÁT Z UWB RADARU

Ako referenčné výsledky pre simulácie na overenie správnosti implementácie predspracovania sme použili program Matlab. Na Obr. 5 a Obr. 6 sú zobrazené výsledky predspracovania dát. Vzhľadom na to, že sa použila metóda paralelného predspracovania, kde sa použili dve postupnosti sú tieto výsledky prezentované ako komplexné číslo, kde reálna zložka obsahuje výsledky prvej postupnosti dát a imaginárna zložka druhej postupnosti dát.

Priebehy na Obr. 5 znázorňujú predspracovanie prvej postupnosti dát, kde prvý graf ukazuje výsledok z programu Matlab, druhý graf ukazuje výsledok z programu napísaného v jazyku C s použitím FFTW knižnice pre rýchlu Fourierovu transformáciu a tretí graf znázorňuje chybu výpočtu, modrý priebeh je výsledný referenčný priebeh vyjadrený v logaritmickej mierke a červený priebeh je chyba získaná ako rozdiel v logaritmickej mierke. Odstup chybového od referenčného priebehu určuje veľkosť chyby výpočtu. Chyba vyjadrená ako súčet kvadrátov pre prvú postupnosť je  $2,211 \cdot 10^{-7}$ , rozdiel referenčného signálu od chybového je  $9,539 \cdot 10^{-10}$ . Na Obr. 6 sú znázornené priebehy pre druhú postupnosť dát, kde grafické výsledky sú vyjadrené rovnakým spôsobom ako v predchádzajúcom prípade. Chyba vyjadrená ako súčet kvadrátov pre druhú postupnosť je  $2,317 \cdot 10^{-7}$ , rozdiel referenčného signálu od chybového je  $9,360 \cdot 10^{-8}$ .



Obr. 5 Výsledky paralelného predspracovania prijatých dát  $x[nT]$  z UWB radaru.



Obr. 6 Výsledky paralelného predspracovania prijatých dát  $y[nT]$  z UWB radaru.

UWB radar posiela dáta každých  $70\text{ ms}$  a preto je dôležité, aby predspracovanie a celé spracovanie dát bolo optimalizované na rýchlosť s čo najmenšou chybou výpočtu. Tab. 1 obsahuje časy jednotlivých funkcií, ktoré je nevyhnutné realizovať v rámci predspracovania dát, položka *celkom preproc()* vyjadruje trvanie celej funkcie predspracovania a v jej súčasťou sú funkcie priamej FFT, násobenia vo frekvenčnej oblasti a spätná FFT.

Tabuľka 1  
 Časy trvania jednotlivých funkcií paralelného predspracovania dvoch postupností dát z UWB radaru

Názov funkcie	Trvanie funkcie
Priama FFT	423,41 $\mu\text{s}$
Komplexné násobenie	93,87 $\mu\text{s}$
Spätná FFT:	429,78 $\mu\text{s}$
celkom preproc():	1,16236 $\text{ms}$



## ZÁVER

Použitie Mini PC MK802 III s programom napísaným v jazyku C sa ukázalo ako veľmi vhodná alternatíva za programy Matlab a LabVIEW predovšetkým vo vložených aplikáciach, kde cena a spotreba sú kritickými faktormi. Použitie optimalizovanej FFTW knižnice na rýchlu Fourierovu transformáciu umožnilo dosiahnuť veľkú rýchlosť predspracovania dát s trvaním 1,16236 *ms* a s dostatočne malou chybou výpočtu, kde v prípade obidvoch postupností dát je odstup chybového signálu od referenčného približne  $-127$  *dB*.

## POĎAKOVANIE

Tento článok bol vytvorený v rámci realizácie projektu UWB senzorové siete krátkého dosahu na detekciu, lokalizáciu a sledovanie pohybujúcich sa osôb (VEGA projekt č 1/0563/13) a realizácie projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] Rikomagic. (2014) Prehľad, mini počítač dual core mk802 iii. [Online]. Available: <http://www.rikomagic.com/en/product/showpro?id=31&pid=19.html>
- [2] ARM. (2014) Procesor arm cortex-a9. [Online]. Available: <http://www.arm.com/products/processors/cortex-a/cortex-a9.php>
- [3] J. Hutira, "Vložené aplikácie na báze mk802 III," Diplomová práca, Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2014.
- [4] Picuntu. (2014) Domovská web stránka picuntu. [Online]. Available: <http://ubuntu.g8.net/>
- [5] J. J. Nielsen. (2014) Domovská web stránka mixfft knižnice. [Online]. Available: <http://www.corix.dk/html/fft.html>
- [6] M. Frigo and S. G. Johnson. (2014) Domovská web stránka fftw knižnice. [Online]. Available: <http://www.fftw.org/>
- [7] E. productivity tools. (2014) Fft čistej reálnej postupnosti alebo paralelizácia fft. [Online]. Available: <http://www.engineeringproductivitytools.com/stuff/T0001/PT10.HTM>

# FIR Filter Implementation for eDiViDe Remote Laboratory

<sup>1</sup>Oto Petura, <sup>2</sup>Miloš Drutarovský, <sup>3</sup>Michal Varchola

Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>oto.petura@student.tuke.sk <sup>2</sup>milos.drutarovsky@tuke.sk <sup>3</sup>michal@varchola.com

**Abstract** – Our design’s main objective is to deliver real-life experience with digital FIR filters to students using a remotely accessible station. To pursue this objective we designed a project with the NIOS processor and our FIR filter implementation on the Altera Cyclone III development board. We created this design as part of the eDiViDe project. By access to this design students will have access to the reconfigurable FIR filter implementation suitable for a variety of applications. Since the design is running on real hardware, rather than a simulation environment, students will have to overcome issues connected with hardware limitations. The remote access to the hardware will create good real-life experience, especially for users who can’t access it normally.

**Keywords** – eDiViDe, FIR filter, Altera, FPGA, remote lab

## I. EDIVIDE PROJECT

The eDiViDe project [1] aims at creating and maintaining a network of remotely accessible FPGA stations shown in Fig. 1. This can provide students, who do not have access to the hardware directly, with an opportunity to test their designs on real hardware, using remote access. Remote access is provided using a unified web interface which enables students to check the device status and control the device using the provided controls.

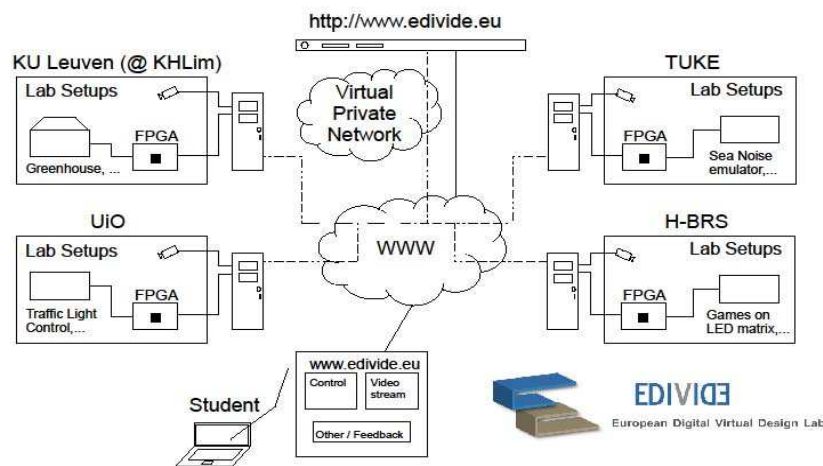


Fig. 1: eDiViDe platform overview [2].

Our FIR filter design is running on an Altera Cyclone III FPGA board [3]. The board is connected to the linux server, which provides remote access through the internet.

## II. DESIGN UTILIZING THE FIR FILTER

Our design consists of 4 major blocks, depicted in Fig. 2. The NIOS II processor [4] is responsible for all communication with the outer world. It manages UART communication for commands, data input/output, and communication with audio codec for audio filtering.

RAM memories for coefficients and samples are dual port RAM memories, which are accessible from NIOS processor and also from FIR filter logic. The coefficients memory is used for

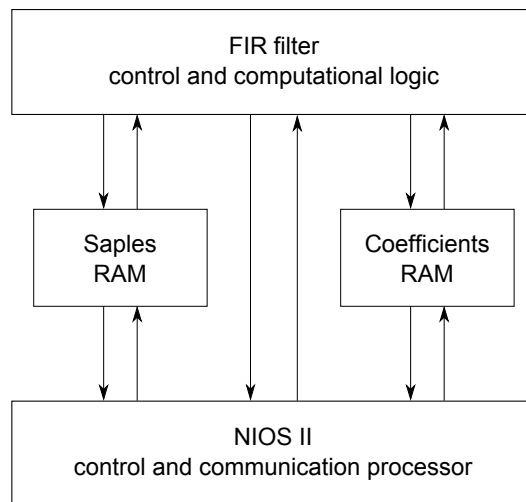


Fig. 2: Block diagram of the FIR filter implementation.

coefficients storage and has a capacity of 16384 bits, which is 1024 filter coefficients. Samples buffer is created by identical memory block as coefficients memory, but is used for temporary sample storage, as a delay line of FIR filter.

FIR filter logic consists of finite state machine, which is responsible for retrieving an appropriate samples and coefficients and computation of the filter.

### III. FIR FILTER IMPLEMENTATION

FIR filter in this design uses 16-bit fixed point arithmetics using fractional 1.15 format [5]. It is implemented in recursive form shown in Fig. 3, which means, that it multiplies always only 2 numbers and then accumulates the result. Multiply And Accumulate (MAC) operation uses 40-bit accumulator, from which 16 output bits are drawn at the end of filtering phase.

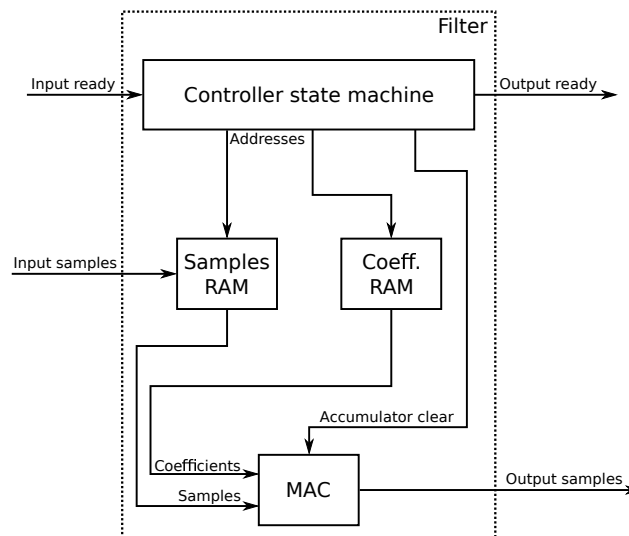


Fig. 3: Internal block diagram of implemented FIR filter.

FIR filter logic uses coefficients and samples memories for data input and one control starting signal. Output is written to 16-bit output register and filter logic uses also output ready signal for indication of valid output data.

If filter receives start signal, it sequentially computes the intermediate values by loading coefficients and appropriate samples from memories and multiplying them together. MAC function automatically accumulates results. After all filter coefficients are used, the logic loads output value into the output register and outputs output ready signal. In the Fig. 4 the state diagram of the finite state machine is shown.

The reset signal is asynchronous and immediately puts the filter to a RESET state. In this state all internal registers are reinitialized with zeros.

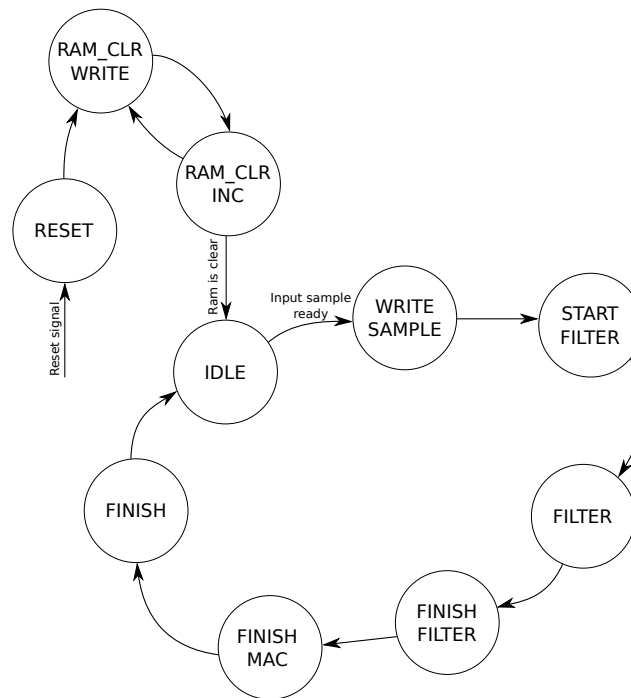


Fig. 4: State diagram of the filter finite state machine.

State machine then goes to RAM\_CLR\_WRITE state, which writes zero value to the sample memory. In RAM\_CLR\_INC state the memory address counter is incremented to erase next address. After whole memory is erased state machine goes into IDLE state.

In IDLE state it waits for the input sample ready signal. After receiving the signal it goes into the WRITE\_SAMPLE state.

In the WRITE\_SAMPLE state, the state machine writes the input sample into the samples memory.

After the sample is written, the filtering process can start. The initialization phase of the filter is done during the START\_FILTER state. During this phase, the state machine waits for the memory to finish writing of the sample.

In the FILTER state, the filtering is performed. State machine sweeps through both memories and multiplies samples with appropriate coefficients. After it reaches the end of the memories, it changes the state to FINISH\_FILTER.

The FINISH\_FILTER state compensates for the delay of the output registers of the memories. Similarly the FINISH\_MAC state compensates for the delay of the MAC unit registers.

After FINISH\_MAC state, all samples are processed and the accumulator holds the output value. The state machine goes into the FINISH state, when it puts the 16-bit output value from the accumulator to the output register and triggers the output sample enable signal.

Filter is equipped with memories, that can hold 1024 coefficients and samples, so 1024 is the maximum number of coefficients for our filter implementation.

The overhead of the filtering process takes 11 clock cycles. These include delays at the input registers of each block of the system and also output registers of memories and MAC unit.

The filtering process takes 1024 clock cycles, one coefficient at a time. The 11 clock cycles are the overhead, so filtering of one sample takes 1035 cycles. The maximum sampling frequency of the input signal can be computed as follows:

$$f_{smax} = \frac{f_{clock}}{N_{cycles}} \quad (1)$$

where  $f_{clock}$  is the frequency of the clock and  $N_{cycles}$  is a number of clock cycles it takes to compute the operation.

Our design is running at 50 MHz clock frequency, so using the (1) the maximum sampling frequency for our design is:

$$f_{smax} = \frac{f_{clock}}{N_{cycles}} = \frac{50 \times 10^6}{1035} = 48309 \text{ Hz} \quad (2)$$

The (1) shows, that maximum sampling frequency of the input signal for our design is 48.3 kHz, which means, that our design can effectively filter input signal with sampling frequency of 48 kHz, which makes it possible to use for audio filtering.

#### IV. NIOS II PROCESSOR

Since our project is supposed to be part of eDiViDe network of remote accessible FPGA stations and therefore it has to communicate with the web interface using specified protocol, we decided to add NIOS II softcore processor to provide flexible support for these needs. NIOS processor is responsible for the communication with outer world and data routing to and from the FIR filter. The Fig. 5 shows the block diagram of connections between NIOS processor and other embedded peripherals.

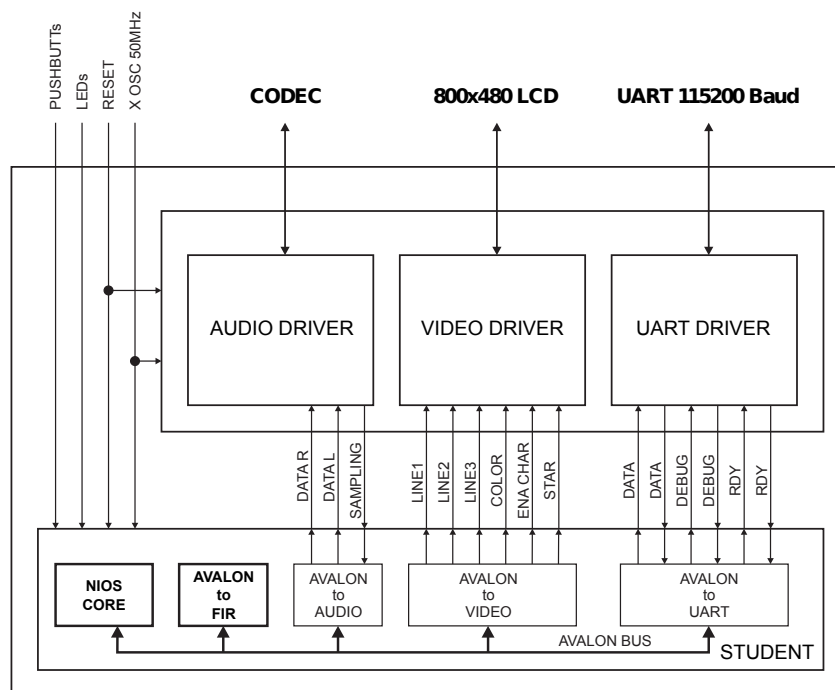


Fig. 5: Block diagram of NIOS processor and peripherals interconnections.

As shown in the Fig. 5, the processor system is responsible for the data management of the entire design. All data from input peripherals (UART, audio) are going through NIOS processor, which then routes the data to internal (FIR) or output peripherals (UART, audio). We included also video driver support for the graphical screen on the board, but in current state we do not use the video in this design. We included the video driver in this design, because we intend to use it for providing graphical feedback to the user, e.g. FFT of the input and output signals.

The UART peripheral is used for the communication with the server, which then forwards transmitted data to the web interface. UART is used mostly for controlling the design through web interface and changing the operation modes. Our design has several operation modes.

##### Audio loopback

Audio samples from line input are directly sent back to audio codec output. This allows user to listen to unfiltered sound.

##### Audio filtering

Audio samples from line input are filtered and then sent to the audio output. The users hears filtered sound.

##### Raw data filtering

This mode allows user to upload the initial state of the filters delay line and also data samples to be filtered. After the data are uploaded and filtered the system will send filtered data samples back to the user. This way the user has complete knowledge and control of the initial state of the filter and the data filtered.

NIOS processor is also used to set the filter coefficients. This is done during the initialization state of the system and coefficients are loaded from the C header file, user has to provide before accessing the design. This approach allows user to test various filter designs.

## V. EXPERIMENTAL RESULTS

During the testing of our project, we encountered a problem, where the signal, that came from the board was not as expected, even if we did not use the filter. The signal seemed to be filtered even if the device was in audio loopback mode. The study of the setup and connections between the board and the server, we found out, that the problem is an external sound card used on the server. There is no internal (on-board) sound device on the server, so to provide sound support we need to use an external device. We tested the external sound card on one of our laptops, which was equipped also with internal card. We played the white noise through the output interfaces and recorded it with input interfaces of both cards. Fig. 6 shows the FFT spectrum of recorded signals.

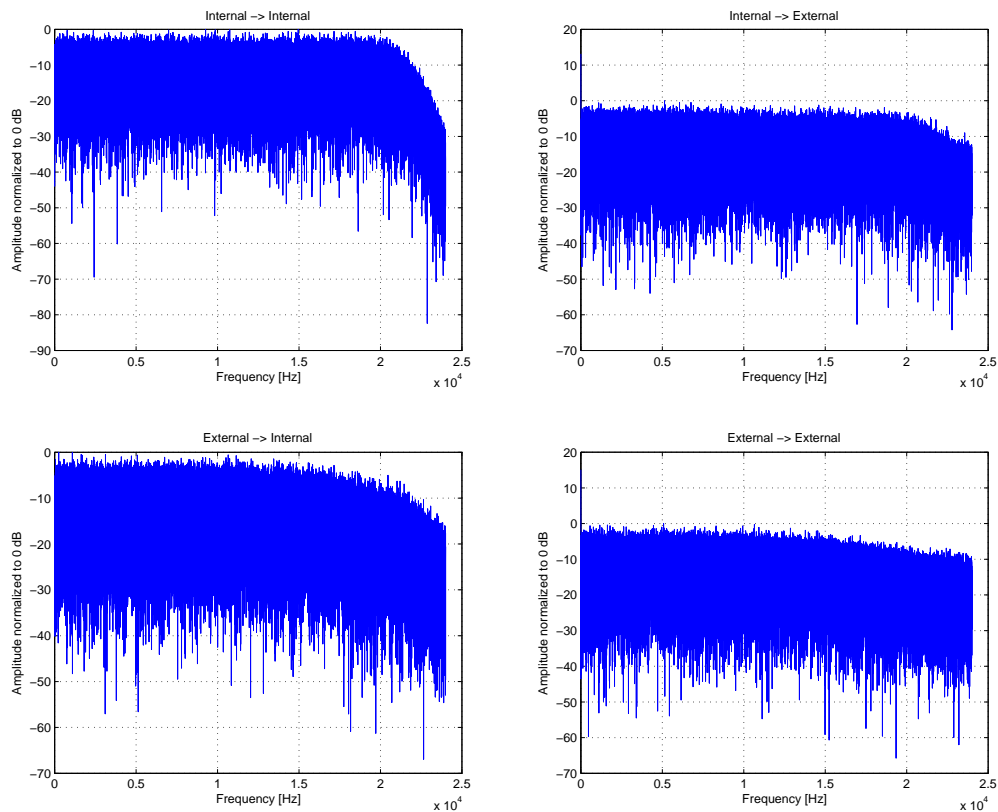


Fig. 6: Characteristics of outputs and inputs of the sound cards.

The graphs in Fig. 6 show, that the external sound card has an output filter with much lower cutoff frequency than the internal one. This affects every signal, that is reproduced by the external sound card on the server. The input of the external card also does not have a filter which can suppress DC component of the input signal.

The issues presented here are connected with the sound card used and can be suppressed only by chose of the hardware used. This type of problem is hard to foresee, because characteristics of the input and output filters of the sound cards are not always public. Problems like this are also strongly dependant on the hardware and do not occur in simulations. That is one of the reasons, why hardware testing is essential in the development process and in education.

### ACKNOWLEDGMENT

This work was supported by European Digital Virtual DEsign (EDIVIDE) lab project No.518565-LLP-1-2011-1-BEERASMUS-ESMO funded by LLP Erasmus

### REFERENCES

- [1] [Online]. Available: <http://www.edivide.eu/>
- [2] J. Vandorpe, J. Vliegen, R. Smeets, N. Mentens, M. Drutarovsky, M. Varchola, K. Lemke-Rust, P. Plöger, P. Samarin, D. Koch, Y. Hafting, J. Tørresen, "REMOTE FPGA DESIGN THROUGH EDIVIDE – EUROPEAN DIGITAL VIRTUAL DESIGN LAB."
- [3] Altera. Nios II Embedded Evaluation Kit, Cyclone III Edition. [Online]. Available: <http://www.altera.com/products/devkits/altera/kit-cyc3-embedded.html>
- [4] ——. Nios II Processor: The World's Most Versatile Embedded Processor. [Online]. Available: <http://www.altera.com/devices/processor/nios2/ni2-index.html>
- [5] Numeric Formats. [Online]. Available: <http://phobos.i.et.unipi.it/~pieri/EdT2/2181Manual/AppendixC.pdf>

# A Low Dropout P-Channel MOSFET based GPS Backup Circuit

<sup>1</sup>Martin RIŠKO, <sup>2</sup>Miloš DRUTAROVSKÝ

<sup>1</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>martin.risko.2@student.tuke.sk, <sup>2</sup>milos.drutarovsky@tuke.sk,

**Abstract** — This article presents a special way to use a P-channel MOSFET transistor as a switch in efficient low dropout backup circuitry of Quectel GPS module. The described transistor configuration was used in a master's thesis circuit design allowing for the connection and disconnection of a power supply voltage from an embedded system and low dropout backup of GPS module from a supercapacitor.

**Keywords** — MOSFET, switch P-Channel, Quectel GPS module, supercapacitor

## I. INTRODUCTION

MOSFET transistors are mostly used for signal amplification or switching applications. In the diploma thesis [1], a special electronic switch was designed (featuring a P-channel MOSFET transistor) for connecting and disconnecting a GPS module from its power source (a supercapacitor in this case) in the event of a board supply voltage removal. The designed MOSFET switch circuit was derived from a similar diode variant switch circuit suggested by the respective GPS manufacturer data sheet [2]. This article details the modifications made.

## II. ORIGINAL SWITCHING CIRCUIT

### A. The Quectel L50

The Quectel L50 GPS module is a low cost GNSS module available for under €10. It offers an input voltage range of 1.71 V to 1.89 V, a current consumption during acquisition of 33 mA, and a hibernation current draw of 14  $\mu$ A. With a hot start under one second, assisted GPS capability, a noteworthy tracking sensitivity of -163 dBm, special *fast time-sync* modes of operation, the L50 proves to be quite appealing apart for its poor documentation and support.

### B. Quectel L50 Power Supply

The Quectel L50 GPS module requires two individual power supplies for nominal operation. These input voltages are applied at two pins; the *VIO/RTC* pin, and the *VCC* pin. The *VIO/RTC* voltage supply supplies power to the internal real time clock, the volatile RAM memory, and other internal voltage critical circuitry. The *VCC* pin supplies power to the internal ARM processor, and other power-thirsty blocks.

While functioning, the L50 cycles through three different modes of operation; the acquisition mode, tracking mode, and the hibernation mode. During the acquisition mode, the L50 searches for satellites. During this mode of operation, the GPS module has the highest current consumption. During the tracking mode, the current consumption is lower, as the L50 simply tracks the satellites. These current consumptions are in the order of tens of milliamperes. During the hibernation mode however, the L50 draws a current in the order of tens of microamperes.

The transitioning between these modes of operation is where the two voltage supplies of the L50 come into play. Because the acquisition and tracking modes of operation require an active processor core, the L50 requires 1.8 V nominal voltages at both the *VCC* and *VIO/RTC* pins. During the hibernation mode of operation, the L50 only keeps track of its stored data, and remains in a low power mode with the main processor core and supporting circuits off. This allows for the lower current consumption. During this mode of operation, active power is only required at the *VIO/RTC* pin.

To sum it up, for continuous operation, the L50 must always have a nominal voltage supply at its VIO/RTC pin, and only requires a nominal voltage supply at the VCC pin during the acquisition and tracking modes of operation.

Because the L50 works with low current consumptions, it is ideal for low-power applications. While being powered from a renewable energy source like a photovoltaic cell (board voltage source-BVS), which outputs varying energy values, the L50 requires an energy reservoir to provide energy to the VIO/RTC pin at all times (even when the BVS drops to zero volts). Figure 1 shows a fail-safe circuit backing up the VIO/RTC pin in the event of an inactive BVS. This circuit utilizes a small battery as a backup energy reservoir and uses silicon diodes for rectification. The circuit was suggested by Quectel, the manufacturer of the L50.

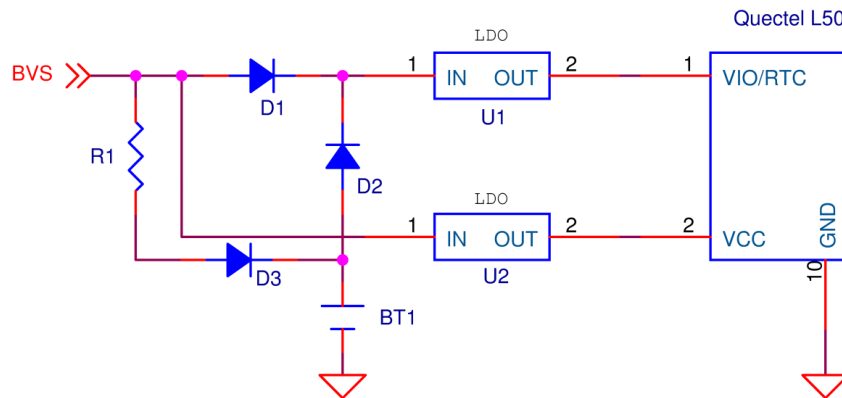


Fig. 1 The L50 power supply fail-safe circuitry as suggested by Quectel. LDO- Low drop-out regulator (both U1 and U2 have a rated output of 1.8 V), BT1 – Battery, BVS – Board Voltage Supply (ex.: photovoltaic cell).

In figure 1, in the event of an active BVS, LDO U1 will have an input voltage (LDO U1 pin 1) equal to the BVS reduced by the voltage drop of D1 (about 0.6 V). LDO U1 output will be the nominal 1.8 V which will be fed to the VIO/RTC input of the L50. Battery BT1 will be charged via diode D3 and resistor R1. LDO U2 will have an input voltage (LDO U2 pin 1) equal to that of the BVS, and an output voltage equal to the nominal 1.8 V which will supply the VCC pin. In the event of an inactive BVS, the diode configuration will cause the VCC pin to observe a zero voltage, while the VIO/RTC pin remains powered from battery BT1, via diode D2, and LDO U1. In other words, in the event of a power outage at BVS, the VIO/RTC pin voltage will stay active.

### III. NEW SWITCHING CIRCUIT REQUIREMENTS

In the mentioned diploma thesis [1], for experimental reasons, battery BT1 from figure 1 was replaced with a supercapacitor. The utilized microprocessor required a supply voltage of 3.3 V, which meant that after the BVS marker in figure 1, an LDO with a nominal output of 3.3 V was inserted the output of which was not only fed to the LDOs U1 and U2 in figure 1, but also to the voltage supply of the microprocessor.

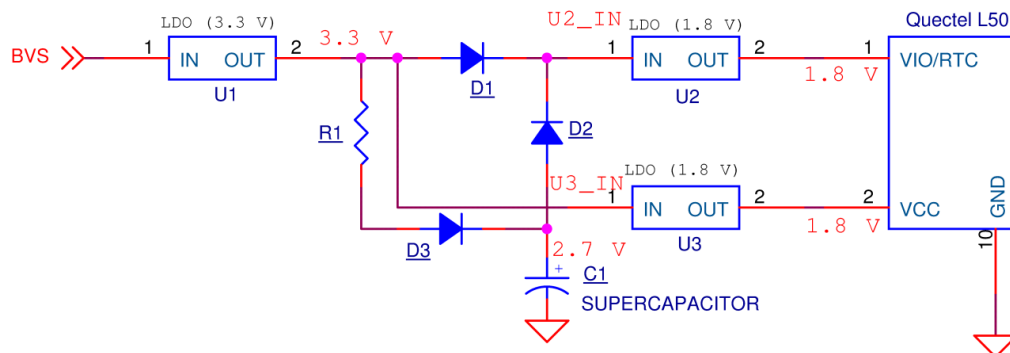


Fig. 2 A slight modification to the rectification circuit suggested by Quectel in figure 1, featuring the supercapacitor, and an additional LDO. The diodes make also make sure the supercapacitor does not discharge into other circuits when BVS drops to zero.



The 3.3 V output of the newly introduced LDO (LDO U1 in figure 2) meant that the supercapacitor in place of battery BT1 could only be charged to a voltage of 3.3 V. Since the minimum voltage requirement for LDOs supplying power to the L50 is 1.8 V (to maintain a 1.8 V output), the supercapacitor would have an operating range of 1.8 V to 3.3 V. If the same diode configuration was to be used as in figure 1, the effective operating range would be reduced greatly due to the voltage drops on the diodes, namely 0.6 V on each diode. The resultant operating range of the supercapacitor would be 2.4 V to 2.7 V. Therefore, a new rectification circuit had to be introduced in place of the wasteful diodes.

#### IV. A P-CHANNEL MOSFET SWITCH

Using a MOSFET transistor as a switch proved to be tricky initially, since the parasitic drain-source diode seemed to always be in the way. But an interesting connection typically used for reverse voltage or battery protection [3,4] was eventually used.

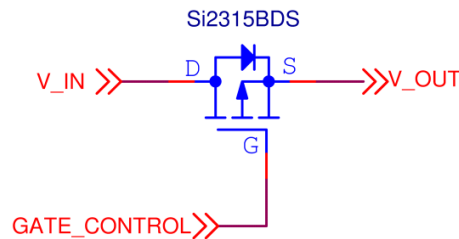


Fig. 3 A P-channel MOSFET being used as a switch

Figure 3 shows the chosen P-channel MOSFET (Si2315BDS) [5] in a switching configuration. *GATE\_CONTROL* is a signal sent from a driving circuit (like a microprocessor for example) for opening and closing the transistor. If the *GATE\_CONTROL* signal is HIGH, current will always be allowed to flow from the drain to the source of the transistor via the parasitic diode, as long as  $V_{IN}$  is greater than  $V_{OUT}$  by at least the diode voltage drop of 0.6 V. In the event of a *GATE\_CONTROL* signal which fully opens the transistor (a LOW signal for a P-channel MOSFET), the parasitic diode is bypassed, and  $V_{OUT}$  will equal  $V_{IN}$ . The advantage of the switching circuit in figure 3 is that no matter the *GATE\_CONTROL* signal level, current will always be allowed to pass from  $V_{IN}$  to  $V_{OUT}$  whereas reverse current from  $V_{OUT}$  to  $V_{IN}$  may be restricted by proper signals at *GATE\_CONTROL*.

#### V. P-CHANNEL MOSFET SWITCH IMPLEMENTATION IN THE DESIGN

To replace the wasteful diodes depicted in figure 2, the MOSFET switch detailed in figure 3 was utilized.

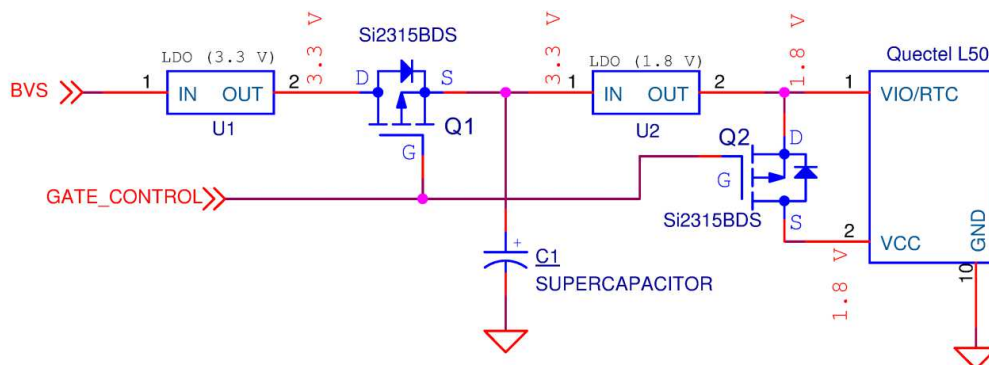


Fig. 4 A modification to the circuit in figure 2. The diode configuration was replaced by two P-channel MOSFETs which also allowed the number of LDOs used to be reduced from three to two.

Figure 4 details the modifications made to the previous LDO and diode configurations. Diodes D1, D2, and D3 were replaced by transistors Q1 and Q2.

Transistor Q1 allows for the supercapacitor to charge during an active BVS voltage and prohibits it from discharging during an inactive BVS voltage. Transistor Q2, when opened, provides a nominal voltage to the VCC pin of the L50. During an inactive BVS, to prevent a fast discharge of the supercapacitor, Q2 closes allowing the GPS module to only draw the respective hibernation mode current. The GATE\_CONTROL signal is provided by an external superior driving circuit like a microprocessor.

The rectification circuit modification also allowed for the reduction in the number of LDOs used. Since transistor Q2 now feeds a 1.8 V voltage from VIO/RTC to VCC, a separate 1.8 V LDO is not needed. This interesting P-channel MOSFET switch configuration actually uses the parasitic diode to its advantage, and offers an optimal switching circuit for the given application.

The GATE\_CONTROL signal in the mentioned diploma thesis [1] is generated by a microprocessor. The microprocessor on the other hand, monitors the BVS voltage via a voltage comparator circuit (using an operational amplifier). In the event of an inactive BVS, the GATE\_CONTROL signal closes the transistors, otherwise, they are left opened.

## VI. CONCLUSION

The presented solution utilizing P-channel MOSFETs for special switching circuits provides a unique way for analog switching and rectification while taking advantage of the transistor's native parasitic diode. The presented solution may be used in a broad range of efficient backup applications.

## ACKNOWLEDGMENT

This work was supported by the project Wireless Sensor Network for wAter QaUality Monitoring (WSN-AQUA) (registration number: HUSK/1101/1.2.1/0091) and the project Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program co-financed from EU funds.

## REFERENCES

- [1] M. Riško, "Embedded Module for Localization and Time Synchronization", Master Thesis, KEMT FEI TU Košice, REV\_A, Košice, 2014, pp. 1–104.
- [2] Quectel, L50 Hardware Design: GPS Module Series, Quectel. 2013-4-12, p. 1–48, [online] [http://www.quectel.com/UploadImage/Downlad/L50\\_Hardware\\_Design\\_V2.0.pdf](http://www.quectel.com/UploadImage/Downlad/L50_Hardware_Design_V2.0.pdf).
- [3] <http://hackaday.com/2011/12/06/reverse-voltage-protection-with-a-p-fet/>
- [4] M. Purschel, "Reverse Battery Protection", Application Note, V 1.0, May 2005, pp.1–6, [online] [https://www.btipnow.com/library/white\\_papers/Reverse\\_Batery\\_Protection.pdf](https://www.btipnow.com/library/white_papers/Reverse_Batery_Protection.pdf).
- [5] "Si2315BSD, P-channel 1.8-V (G-S) MOSFET", data sheet S-80642-Rev.E, Vishay Siliconix, 24-Mar-08, pp.1–9.

# Klasifikácia diskusných príspevkov pomocou metód strojového učenia

*Dávid RIGAS, Kristína MACHOVÁ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

rigas.david@gmail.com, kristina.machova@tuke.sk

**Abstrakt** — Táto práca sa zaoberá klasifikáciou názorov prezentovaných vo filmových recenziách. V práci sú podrobne popísané princípy algoritmov strojového učenia, ich výhody a nevýhody. Taktiež obsahuje definíciu problému analýzy sentimentu a popisuje prístupy k jeho riešeniu. Práca obsahuje návrh metódy analýzy sentimentu so zameraním na určovanie názorovej polarizácie filmových recenzií. Táto metóda bola implementovaná. Vzniknutá aplikácia bola testovaná. Dosiahnuté výsledky sú v príspevku prezentované prostredníctvom grafov pre lepšiu čitateľnosť a zrozumiteľnosť. Práca taktiež analyzuje výsledky, ktoré boli dosiahnuté pomocou Bayesovho naivného klasifikátora. V závere sú popísané možnosti vylepšenia aplikácie.

**Kľúčové slová** — Klasifikácia názorov, Bayesov naivný klasifikátor, trénovacia množina, recenzia

## I. ÚVOD

V súčasnej dobe si ľudia nevedia predstaviť svoj každodenný život bez používania internetu, ktorý využívajú predovšetkým na vyhľadávanie informácií, produktov a komunikáciu. Internet predstavuje aj priestor vyjadrovania svojich názorov prostredníctvom komentárov, recenzií produktov alebo služieb. A práve analýze názorov prezentovaných v rámci sociálneho webu je venovaný tento príspevok.

Rôzne spravodajské portály umožňujú svojim čitateľom vyjadriť svoj názor na konkrétne objekty alebo články iných prispievateľov. Čitatelia môžu viesť medzi sebou diskusie a zhodnotiť prínos a objektivnosť článku od niektorého autora. Existujú aj portály, ktoré slúžia na hodnotenie produktov a obchodov, prostredníctvom ktorých ľudia nakúpili. Tieto svoje skúsenosti s nákupom a tovarom môžu vyjadriť prostredníctvom on-line recenzie. Táto recenzia môže ovplyvniť rozhodovanie potencionálnych zákazníkov o nákupe konkrétneho produktu alebo rozhodovanie o nákupe u konkrétneho dodávateľa produktov alebo služieb. [1]

Práve tieto príspevky a názory ľudí môžu byť pre firmy veľmi užitočné. Sú to veľké množiny dát, ktoré sú nositeľmi názorov. V prípade, ak dokáže firma tieto dáta spracovať, môže sa dostať do konkurenčnej výhody a získa informácie od zákazníkov bez vykonávania marketingových prieskumov. [2]

Strojové učenie obsahuje algoritmy, pomocou ktorých je možné takéto vyhodnocovania realizovať v praxi. Cieľom tejto práce je klasifikácia diskusných príspevkov pomocou metód strojového učenia.

## II. ANALÝZA NÁZOROVEJ POLARITY PRÍSPEVKOV

Bola navrhnutá metóda analýzy názorovej polarizácie príspevkov tvoriacich konverzačný obsah, založená na použití Naivného Bayes klasifikátora. Tento algoritmus bol zvolený pre jeho použiteľnosť v riešení širokého spektra klasifikačných problémov. Cieľom bolo vytvorenie aplikácie, ktorá na základe vytvorenej trénovacej množiny dokáže natrénovať Bayes klasifikátor na klasifikáciu diskusných príspevkov k pozitívnemu alebo negatívnemu názoru. Táto aplikácia bola vytvorená na platforme .NET. Zdrojový kód aplikácie bol napísaný v jazyku C#, ktorý má slúžiť a bol vytvorený pre vytváranie rôznych aplikácií v technológii .NET. Aplikácia bola konkrétne vytvorená v technológii WPF. Táto technológia patrí medzi jednu z novších technológií .NET a má slúžiť ako náhrada pre technológiu Windows Forms. Okrem jazyka C# v technológii WPF je využívaný aj XAML, ktorý je odvodený od jazyka XML. [3]

Ako vstupné dáta pre túto aplikáciu boli použité slovenské komentáre z webovej stránky [www.csfd.cz](http://www.csfd.cz). CSFD je skratka pre Česko-Slovenská filmová databáza a predstavuje najväčšiu on-line databázu filmov a ich hodnotení na území Slovenska a Česka. Na tejto doméne je možné pridávať hodnotenia a komentáre k jednotlivým filmom, ktoré používateľ tejto domény videl. Týmto spôsobom môžu návštevníci tejto domény zdieľať vlastné skúsenosti a postrehy, a na základe komentárov ostatných ľudí sa rozhodnúť, či si daný film pôjdu pozrieť do kina alebo sa rozhodnú pre pozretie iného filmu.

Celkovo bola k dispozícii množina 63 975 rôznych príspevkov od používateľov, ktorí hodnotili jednotlivé filmy a seriály. Táto množina príspevkov pozostávala zo siedmich stĺpcov: ID, Hodnotený film, Nick hodnotiteľa, Kategória hodnoteného filmu, Komentár hodnotiteľa filmu, Dátum hodnotenia, Rating (Ohodnotenie daného filmu). Pre prezentovanú aplikáciu boli podstatné stĺpce, ktoré obsahovali komentár a ohodnotenie videného filmu. Bolo potrebné vyfiltrovať tie filmy, ktoré boli hodnotené ako veľmi dobré. Medzi tieto filmy boli zaradené tie, ktoré :

- Sú ohodnotené štyrmi hviezdčkami (\*\*\*\*)
- Sú ohodnotené piatimi hviezdčkami (\*\*\*\*\*)

Taktiež boli filtrované aj veľmi zlé filmy – podľa používateľov. Medzi tieto filmy boli zaradené tie, ktoré:

- Sú ohodnotené ako odpad (odpad!)
- Sú ohodnotené jednou hviezdčkou (\*)

Ostatné príspevky nachádzajúce sa v celkovej množine príspevkov boli pre túto aplikáciu nepodstatné, keďže sa nezameriavala na neutrálne príspevky ale iba na pozitívne a negatívne príspevky.

### III. VYUŽITIE BAYSOVHO NAIVNÉHO KLASIFIKÁTORA

Aby bolo možné natréňovať Bayesov naivný klasifikátor, v prvom rade bolo potrebné pripraviť tréningovú množinu. Z databázy komentárov k filmom z domény [csfd.cz](http://csfd.cz) bolo vyselektovaných 400 príspevkov. Z toho 100 príspevkov bolo použitých na vytvorenie tréningovej množiny, z ktorých bolo:

- 50 príspevkov hodnotených ako veľmi pozitívne
- 50 príspevkov hodnotených ako veľmi negatívne

Z týchto príspevkov bolo potrebné vybrať tie slová, ktoré boli nositeľmi pozitívneho alebo negatívneho významu, a na základe ktorých by program vedel klasifikovať príspevky do konkrétnej triedy. Na tento účel bol pripravený korpus slov v slovenskom jazyku, ktorý obsahoval 19 403 slov v rôznych gramatických tvaroch. Obsahoval stĺpce: Slovo, Slovný druh, Typ pocitu, Váha slova.

Keďže v aplikácii sa nevyužívajú slovníky slovných druhov ani váhy, ale je vytváraná tréningová množina, bol použitý iba stĺpec so slovami a informácia, či slovo je pozitívne alebo negatívne. Pri vytváraní tréningovej množiny sa ukázalo, že tento korpus neobsahuje všetky slová, ktoré boli nositeľmi názoru a často sa vyskytovali vo filmových recenziách. Na základe tohto zistenia bolo potrebné korpus slov rozšíriť o ďalších 3 756 slov a celkový počet slov v korpuse bol 23 156.

Pre výber slov z vytvoreného korpusu, ktoré sa vyskytujú v komentároch bola vytvorená aplikácia, ktorej úlohou bolo:

- Vyhľadanie kľúčových slov z príspevku
- Určenie, koľkokrát sa slovo v danom príspevku nachádza
- Zistenie, či sa konkrétne slovo nachádza v predošliých príspevkoch zaradených do tréningovej množiny
- Uloženie tých príspevkov, ktoré boli zaradené do tréningovej množiny do samostatného textového súboru
- Uloženie slov vybraných z príspevkov a nachádzajúcich sa v korpuse slov do samostatného textového súboru.

Táto aplikácia poslúžila na vytvorenie tréningovej množiny, ktorá obsahovala:

- 100 komentárov k filmom
- 290 základov slova
- 7 589 rôznych tvarov slov

Z vytvoreného korpusu boli priradené ku každému základu slova rôzne tvary tohto slova. Ďalej bolo potrebné vytvoriť databázu tréningovej množiny, ktorá v každom príspevku obsahuje

hodnoty 1 alebo 0 k jednotlivým kľúčovým slovám:

- 1 znamená, že slovo sa v danom príspevku nachádza
- 0 znamená, že slovo sa v danom príspevku nenachádza

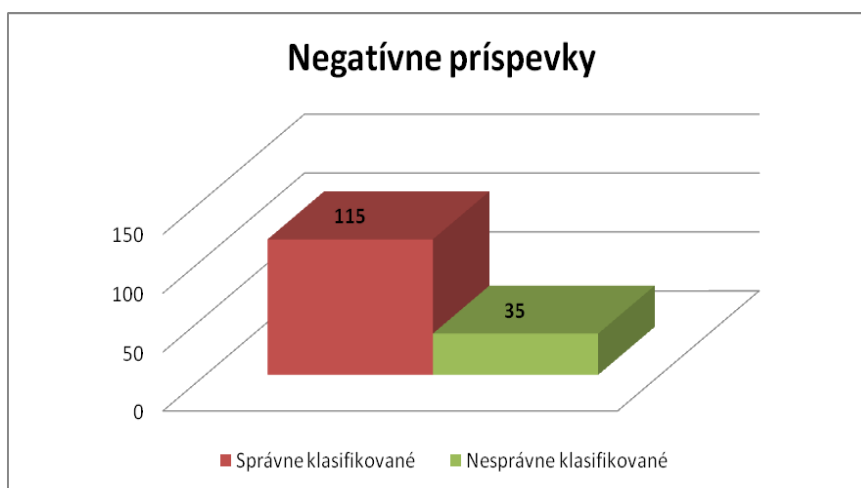
Táto databáza zároveň obsahuje aj stĺpec trieda. Tento stĺpec určuje, či daný príspevok patrí k pozitívnej alebo negatívnej triede:

- Znamienko „+“ znamená, že príspevok z trénovacej množiny je zaradený do pozitívnej triedy
- Znamienko „-“ znamená, že príspevok z trénovacej množiny je zaradený do negatívnej triedy

Na základe tejto databázy boli vyčíslené pozitívne a negatívne hodnoty slov. Tieto hodnoty boli následne použité pre vyčíslenie pravdepodobnosti klasifikácie analyzovaných diskusných príspevkov ku pozitívnemu alebo negatívnemu názoru.

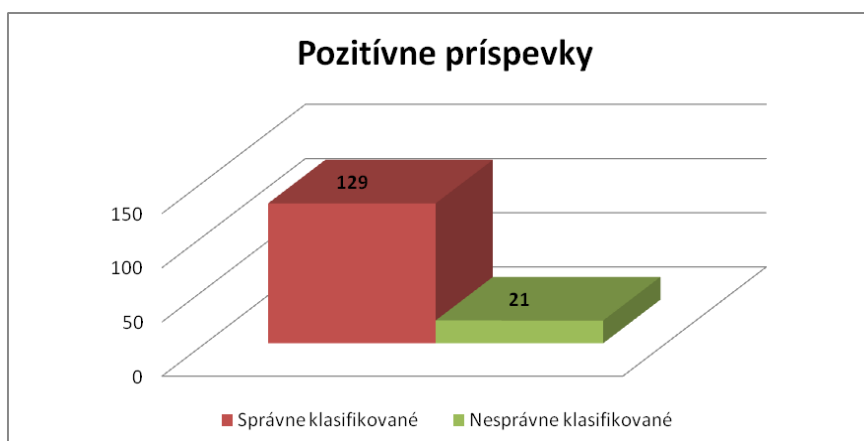
#### IV. TESTOVANIE APLIKÁCIE

Na testovanie aplikácie bolo použitých 300 vybraných recenzií používateľov domény www.csfd.cz, ktorí hodnotili jednotlivé filmy. Z toho bolo 150 pozitívnych a 150 negatívnych. Výsledky týchto testov sú zobrazené na nasledujúcich grafoch. Z celkového počtu 150 negatívnych testovacích príspevkov bolo správne klasifikovaných 115 príspevkov. Nesprávne klasifikovaných bolo 35 príspevkov (viď Obr.1). To znamená, že presnosť klasifikovaných negatívnych príspevkov dosiahla 77%.



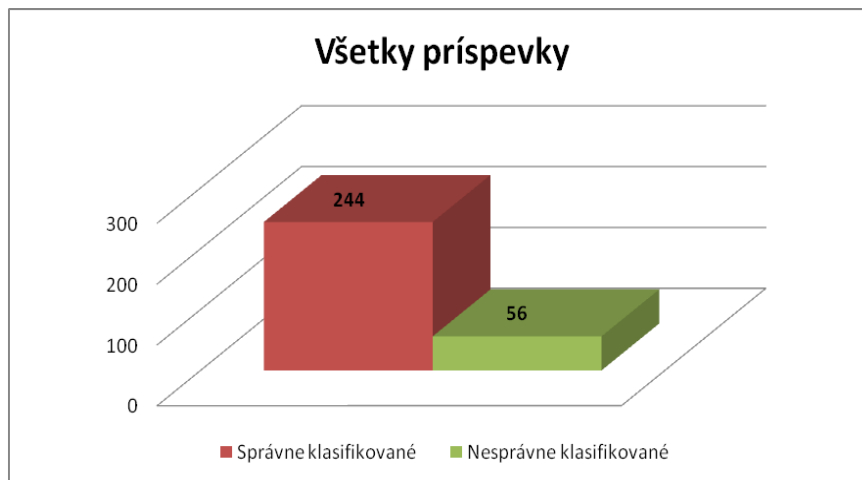
Obr. 1 Hodnotenie efektívnosti klasifikácie na negatívnych príspevkoch

Z celkového počtu 150 pozitívnych vyhodnocovaných testovacích príspevkov bolo správne klasifikovaných 129 príspevkov. Nesprávne klasifikovaných bolo 21 príspevkov (viď Obr.2). To znamená, že presnosť klasifikovaných pozitívnych príspevkov dosiahla 86%.



Obr. 2 Hodnotenie efektívnosti klasifikácie na pozitívnych príspevkoch

Z celkového počtu 300 vyhodnocovaných testovacích príspevkov bolo správne klasifikovaných 244 príspevkov. Nesprávne klasifikovaných bolo 56 príspevkov, čo ilustruje Obr.3. To znamená, že presnosť klasifikovaných negatívnych príspevkov dosiahla 81%.



Obr. 3 Hodnotenie efektívnosti klasifikácie na všetkých príspevkoch

Z celkového počtu 45 vyhodnocovaných testovacích príspevkov, ktoré boli zaradené do kategórie „Odpad“ bolo správne klasifikovaných 39 príspevkov. Nesprávne klasifikovaných bolo 6 príspevkov. To znamená, že presnosť klasifikovaných negatívnych príspevkov dosiahla 87%.

#### V. ZÁVER

Na vytvorenie trénovacej množiny bolo použitých 100 príspevkov. Na základe testov, ktoré boli vykonané prostredníctvom navrhutej aplikácie na 300 filmových recenziách bola dosiahnutá presnosť klasifikácie 81 percent. Navrhnutá aplikácia dosahuje vyššiu presnosť pri klasifikácii pozitívnych recenzií o 9 percent, ako pri klasifikácii negatívnych recenzií.

Aplikácia dosahuje lepšiu presnosť pri recenziách, ktorých hodnotenie filmu je veľmi dobré alebo veľmi zlé. Pri filmoch, ktoré boli ohodnotené ako „Odpad“ bola presnosť vyššia o 15 percent ako u filmov, ktoré boli ohodnotené jednou hviezdíčkou. Pri pozitívnych recenziách tento rozdiel nebol, až taký výrazný.

Pre zvýšenie presnosti klasifikácie názoru u negatívnych recenzií by bolo potrebné rozšíriť trénovaciu množinu o ďalšie príspevky, ktoré by obsahovali ďalšie slová, ktoré majú zápornú polaritu. Ďalšou možnosťou, ako by bolo možné dosiahnuť vyššiu presnosť by bolo zahrnutie aj dvojíc slov, pri ktorých sa mení ich polarita a ich význam. Na základe vykonaných testov sa ukázalo, že je možné dosiahnuť pomerne dobrú presnosť pri rozsiahlej trénovacej množine aj pomocou Bayesovho naivného klasifikátora.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] M. Thekwall, K. Buckley, G. Paltoglou, D. Cai, „*Sentiment Strength Detection in Short InformalText.*“ University of Wolverh. ampton, 24s.
- [2] B. Pang, L. Lee, „*Opinion mining and sentiment analysis.*“ Foundation and Trends in Information Retrieval 2, 2008. 94s.
- [3] CH.Nagel, B. Evien, J. Glynn, M. Skinner, K. Watson, „*C# 2008 Programujeme profesionálne.*“, Brno: Computer Press a.s., 2009. 1112s.

# Identifikácia autorstva dolovaním obsahu v sociálnom webe

*Jozef RUŽBARSKÝ, Kristína MACHOVÁ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

jozef.ruzbarsky2@student.tuke.sk, kristina.machova@tuke.sk

**Abstrakt** — Cieľom príspevku je identifikácia autorstva dolovaním obsahu v sociálnom webe. Zámernom je identifikovať hlavne krátke príspevky v diskusiách na fórach, blogoch a ďalších textoch nachádzajúcich sa v sociálnom webe. Identifikácia autorstva je založená na slovníkovom prístupe – na porovnávaní s preddefinovanými slovníkmi (bežný typ, problémový, spamový, autorita). V článku je prezentovaná windows aplikácia, ktorá slúži pri identifikácii typu autora. Táto aplikácia identifikácie autorstva bola implementovaná v programovacom jazyku C vo vývojovom prostredí dev-c++. V závere je vykonané testovanie danej aplikácie a vyhodnotenie úspešnosti navrhnutého riešenia.

**Kľúčové slová** — Identifikácia autorstva, dolovanie názoru, analýza názoru, typ autora

## I. ÚVOD

Význam identifikácie autorstva je stále dôležitý, keďže je potrebné zisťovať typ prispievateľov a tým zabráňovať umiestneniu príspevkov s nevhodným obsahom v rámci rozličných webových diskusií. Taký konverzačný obsah poburujú ostatných používateľov, ktorí hľadajú potrebné informácie alebo poskytujú svoje odborné vedomosti, skúsenosti iným používateľom.

Navrhnuté riešenie pozostáva z dvoch fáz a to z prípravnej fázy a hlavnej fázy. Prípravná fáza je zameraná na tvorbu slovníkov. Slovníky boli tvorené manuálne – prehľadávaním textov nachádzajúcich sa v <http://www.pcforum.sk/>, ktoré boli vstupom pre ďalšie spracovanie. Následne pri prechádzaní bol určený typ autora každého jednotlivého príspevku a slová, ktoré ho vystihovali boli zaznačené do jednotlivých slovníkov. V hlavnej fáze prebiehalo porovnávanie načítaného textového súboru so slovníkmi jednotlivých typov autorov, ktorého výsledkom bolo priradenie daného textu k jednému typu používateľa.

## II. SOCIÁLNY WEB A KONVERZAČNÝ OBSAH

Sociálny web môžeme charakterizovať ako súbor spoločenských vzťahov, ktoré spájajú ľudí prostredníctvom World Wide Web. Sociálny web zahŕňa web stránky, ako aj softvér na podporu a posilnenie sociálnych interakcií. Sociálne interakcie tvoria základ pre rôzne on-line aktivity, ako sú on-line nakupovanie, vzdelávanie, hranie hier a webové stránky sociálnych sietí. Aktivity ľudí na webe a ich komunikácia zvyšuje informácie o ich sociálnych vzťahov, ktoré sú neustále k dispozícii. [1]

Sociálny web sa rýchlo stáva určitým spôsobom života, čoraz viac ľudí navštevuje sociálne siete viackrát denne a trávi tam viac a viac času. Sociálny web zrovnoprávňuje producentov obsahu s jeho konzumentmi.

Konverzačný obsah predstavujú krátke texty venované nejakému objektu, téme (napr. hodnotenie produktu, politickej situácie, filmu). Tento konverzačný obsah vzniká v rámci diskusných fór, blogov, microblogov, chatov, IRC (Internet Relay Chat) a pod. Všetky tieto platformy produkujú databázy názorov, pocitov, nálad a postojov jednotlivých používateľov webu. Databázy neobsahujú štruktúrované dáta, preto sa ďalej spracovávajú – klasifikácia názorov, analýza názoru. [2]

Klasifikácia názorov je zameraná na subjektívnu stránku textu - postoje, názory a pocity autora. Subjektívne informácie sa vždy na niečo vzťahujú, či to už je osoba, udalosť, tovar a podobne. Nositeľmi subjektívnych postojov v tomto prípade sú slová, ktoré vyjadrujú názor, napríklad sú to prídavné mená, slovesa, podstatné mená, príslovky. [3]

### III. IDENTIFIKÁCIA TYPU AUTORA

Identifikácia autorstva je založená na slovníkovom prístupe – ide o porovnávanie načítaného textového súboru s preddefinovanými slovníkmi (bežný typ, problémový, spamový, autorita). Pre tieto potreby bola navrhnutá Windows aplikácia, ktorá je ľahko ovládaná a je prístupná pre všetkých používateľov. V neposlednom rade bolo potrebné vytvoriť aj slovníky jednotlivých typov autorov – tieto boli vytvorené z existujúcich príspevkov nachádzajúcich sa v počítačových fórach.

#### A. Tvorba slovníkov

Tvorba slovníkov bola zložená na manuálnom princípe, čo prispelo k ich kvalite. Boli prejdené jednotlivé diskusné príspevky a bolo určené, aký názor daný príspevok predstavuje. Boli zaznamenané kľúčové slová, ktoré boli považované za časté vzory slov vystihujúcich daný názor a vyskytujúcich sa vo viacerých príspevkoch rovnakého typu názoru. Ďalej boli tieto kľúčové slová zaradené do slovníka toho typu používateľa, ktorý vystihovali.

Jednotlivé slovníky typov používateľov obsahujú v rozsahu približne 50 a viac slov s rôznymi odvodzeniami, skratkami, v rôznych pádoch, vzoroch a taktiež aj slangové podoby slov. Slová, ktoré sa nachádzajú v slovníku sú zadávané len malými písmenami. Zmena veľkých písmen prináša výhodu, lebo slovníky potom nemusia obsahovať všetky tvary jednotlivých slov. Rozsahom sú menšie a tým sa aj celkový čas porovnávania a spracovania skraca. Taktiež to prispieva k vyššej presnosti navrhnutého riešenia na identifikáciu typu prispievateľov. V dnešnej dobe veľa ľudí navštevuje sociálne siete, hlavne rôzne blogy kde je zvykom používať rôzne skratky a slangové slová. Tieto slová používatelia neskôr používajú pri bežných rozhovoroch ako aj pri vyjadrovaní svojich názorov v diskusných skupinách.

#### B. Identifikácia autorstva

Určenie presného typu prispievateľa sa deje na základe zhody načítaného textového súboru v porovnaní s jednotlivými slovníkmi stanovených typov autorov. Podľa každého slovníka je zistená určitá presnosť so zadaným textovým súborom. Následne sa uplatňuje súťaživý princíp, ktorý slúži na výber konkrétneho typu autora a ten je napokon vypísaný ako identifikovaný typ autora.

#### C. Implementácia riešenia

Metóda identifikácie autorstva dolovaním obsahu v sociálnom webe bola implementovaná v programovacom jazyku C vo vývojovom prostredí dev-c++. Výstupom je Windows aplikácia, ktorá je prístupná pre každého používateľa bez ohľadu na to, či má nejaké poznatky o programovaní, respektíve bez ohľadu na to aký používa operačný systém. Aplikácia je spustiteľná na každom počítači.

Použitie aplikácie a získanie výsledku sa skladá z nasledujúcich krokov:

- načítanie textového súboru,
- priebeh funkcií (predspracovanie, porovnávanie so slovníkmi, zaznačenie zhody, výpočet percentuálnej zhody),
- zvolenie požadovaného výsledku,
- prehľad jednotlivých zhôd v slovníkoch.

### IV. TESTOVANIE APLIKÁCIE

Na vyhodnotenie výstupnej aplikácie bola vytvorená testovacia množina, na ktorej bola skúmaná efektívnosť aplikácie. Táto testovacia množina obsahovala 130 príspevkov od rôznych autorov, ktorí vyjadrovali rôzne názory. Boli vypočítané hodnoty presnosti a návratnosti, aby ukázali, či daná aplikácia spĺňa požadované parametre, a či je prospešná pre jej používateľov. V tabuľkách Tab.1 a Tab.2 sú zhrnuté získané výsledky presnosti a návratnosti jednotlivých typov autorov.

Z uvedených výsledkov (viď Tabuľka 1 a Tabuľka 2) vyplýva, že len zopár analyzovaných príspevkov sa nepodarilo jednoznačne určiť. Je možné to pripísať hlavne výskytu textov takého charakteru, že sa v ňom vyskytujú slova viacerých typov. To zabránilo jednoznačnému určeniu typu autora, pretože zhoda v jednotlivých slovníkoch bola rovnaká.

Najlepšie z posudzovaných typov autorov, ktorí boli zvolení vyšiel typ autor – „Problémový“ (hlavne kvôli malému počtu príspevkov daného typu), ale aj „Typ autorita“. Podľa výsledkov je možné povedať, že aplikácia je prospešná a spĺňa požadované parametre. Taktiež je možné považovať danú aplikáciu za užitočnú a použiteľnú na spracovanie textov z rozličných fór na sociálnom webe. Keďže typovanie autorov vychádzala z príspevkov vo webových diskusiách, mohlo by byť nápomocné pri odstraňovaní nevhodných príspevkov okrem toho, že dokáže identifikovať typ prispievateľa.



Tabuľka 1  
Dosiiahnuté výsledky presnosti

Typ autora - prispievateľa	Výsledná hodnota presnosti
Bežný typ	0,8095
Typ autorita	0,875
Spamový typ	0,875
Problémový typ	1
<b>Celková presnosť</b>	<b>0,89</b>

Tabuľka 2  
Dosiiahnuté výsledky návratnosti

Typ autora - prispievateľa	Výsledná hodnota návratnosti
Bežný typ	0,9855
Typ autorita	0,6364
Spamový typ	0,875
Problémový typ	0,7143
<b>Celková návratnosť</b>	<b>0,8028</b>

## V. ZÁVER

V súčasnej dobe čoraz viac ľudí prispieva do rozličných sociálnych sietí a webových diskusií v rámci sociálneho webu svojimi príspevkami, ktoré tvoria konverzačný obsah. Keďže prispievatelia sú rôzneho typu, je potrebné ich odlišiť, a na to poslúži aplikácia identifikácie autorstva, inak povedané identifikácia prispievateľa sociálneho webu.

Tento príspevok bol zameraný na identifikáciu autorstva pomocou použitia slovníkov preddefinovaných skupín. Slovníky rozličných skupín autorov, boli tvorené manuálnym spôsobom, čo prispelo k lepšej presnosti a taktiež k lepšiemu pochopeniu štýlu vyjadrovania sa jednotlivých autorov, ktorí prispievajú do sociálneho webu.

Navrhnutý prístup k identifikácii autorstva prispieva k obmedzeniu príspevkov, ktoré len poburujú používateľov, ktorí hľadajú prospešné informácie. Na základe navrhnutého spôsobu tvorby slovníkov, ktorý vychádza z existujúcich príspevkov v rôznych webových diskusiách sa dosiahne prispôbenie obsahu slovníkov živej reči diskutujúcich. Taktiež je tento prístup nápomocný fóram, blogom a rôznym webovým diskusiám pri obmedzovaní neprospešných príspevkov a pri odhaľovaní autorov, ktorí narušujú diskusie a doterajšie prístupy ich činnosť zatiaľ neodhalili.

Na základe prieskumu a výsledkov experimentov by bolo do budúca vhodné zvýšiť ešte viac presnosť. Presnosť by bolo možné zvýšiť načítaním diskusie, ktoré by obsahovali viaceré príspevky od toho istého autora. Väčšie množstvo príspevkov od toho istého autora pomôže predísť tomu, aby autor bol považovaný za problémového alebo spamového používateľa, len na základe jedného príspevku. Keďže aj príspevky od autorov typu autorita a bežný typ môžu niekedy obsahovať slová, ktoré by mohli navodiť podozrenie, že ide o iný typ autora, nie takého za akého bol doteraz považovaný.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] „*Social web*.“ [Online] [Dátum 30. 04. 2014]. Dostupné na internete: <[http://en.wikipedia.org/wiki/Social\\_web](http://en.wikipedia.org/wiki/Social_web)>.
- [2] „*Sentiment analysis*.“ [Online] [Dátum 30. 04. 2014]. Dostupné na internete: <[http://en.wikipedia.org/wiki/Sentiment\\_analysis](http://en.wikipedia.org/wiki/Sentiment_analysis)>.
- [3] K. Machová, „*Sociálny web, analýza názorov, klasifikácia názorov, dolovanie konverzácií*.“ [Online] Technická univerzita v Košiciach, 2013. [Dátum: 03. 03 2014.] Dostupné na internete: <<http://people.tuke.sk/kristina.machova/prezentacieSaSW/>>

# PWM Modulator for Multilevel Inverter

<sup>1</sup>Marek PÁSTOR, <sup>2</sup>Jaroslav DUDRIK, <sup>3</sup>Viktor ŠLAPÁK, <sup>4</sup>Michal PAJKOŠ

Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>marek.pastor@tuke.sk, <sup>2</sup>jaroslav.dudrik@tuke.sk, <sup>3</sup>viktor.slapak@tuke.sk,  
<sup>4</sup>michal.pajkos@tuke.sk

**Abstract** — The paper describes development of the DSP based multilevel PWM modulator for the laboratory model of the 15-level cascade H-bridge inverter. The switching frequency of the cascade inverter is set to 5 kHz. Design steps together with laboratory measurement are presented.

**Keywords** — DSP, multilevel inverter, sinusoidal PWM

## I. INTRODUCTION

Multilevel inverters have several advantages over classical two-level inverters. These advantages include staircase output voltage, lower du/dt, lower switching frequency, lower THD. When used in grid-connected applications the multilevel inverters produce currents with lower ripple and require smaller filters to attenuate higher-order harmonics.

There are several control techniques for multilevel inverters. Some of them use constant switching frequency and some use variable switching frequency. Pulse Width Modulation technique use constant switching frequency and is often used in multilevel inverters [1,2]. Many papers deal with various modifications of the multilevel PWM modulation [1,3,4]. There are several variations of the PWM in multilevel inverters mainly due to poor level utilization for small modulation indices[5,6]. However, with grid-connected inverter there is no such problem. The inverter output voltage can not be lower than the grid voltage at any time instant.

Digital Signal Processors are often used to control power converters. DPSs often have PWM module. However the PWM module is suited for classical two-level inverters and can not be used for multilevel inverter.

The paper describes development of the DSP based multilevel PWM modulator for 15- level cascade H-bridge inverter. The cascade H-bridge inverter consists of three separated H-bridge inverters and thank to the modified control strategy is capable of creating 15 levels of the output voltage.

## II. MULTILEVEL INVERTER

### A. Topology

The one-phase grid-connected cascade H-bridge inverter (Fig. 1) consists of three H-bridge inverters isolated on the DC side. The cascade inverter is connected to the grid through an LCL filter. The LCL filter is designed for the switching frequency of 5 kHz and the apparent power of 1 kVA. The cascade inverter with d equal DC voltages is capable of creating n voltage levels:

$$n = 2d + 1 \quad (1)$$

The cascade inverter shown in Fig. 1 is thus able to create 7 voltage levels. However, by using uneven DC supply voltages where the DC voltage of the next H-bridge inverter is two times higher than the DC voltage of the previous H-bridge inverter it is possible to increase the number of the output voltage levels to:

$$n = 2^{d+1} - 1 \tag{2}$$

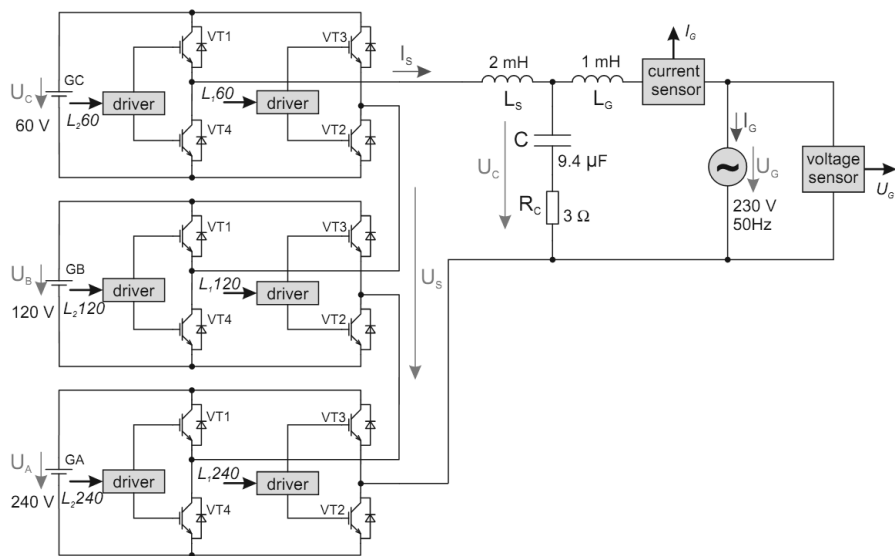


Fig. 1 Grid-connected cascade H-bridge inverter

**B. Multilevel inverters modulation techniques**

There are several techniques available to control the multilevel inverter (Fig. 2). Sinusoidal PWM belongs to high frequency control techniques. Based on carrier signals it can be further divided into Symmetric Disposition, Opposition Disposition and Alternate Opposition Disposition PWM. [7,8]

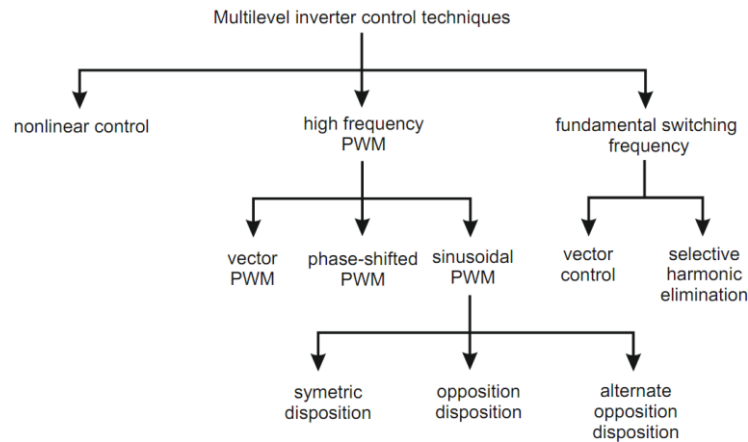


Fig. 2 Multilevel inverter control techniques [7,8]

**C. Sinusoidal PWM**

Pulse Width Modulation (PWM) is a modulation technique that controls the width of the pulse based on the modulation signal. The modulation signal is sinusoidal. There are two quantities defining properties of the multilevel PWM, namely the amplitude ( $m_a$ ) and frequency ( $m_f$ ) modulation indices.

$$m_a = \frac{A_m}{(n-1)A_c} \tag{3}$$

where:  $A_m$  is the modulation signal amplitude,  $A_c$  is the carrier signal amplitude and  $n$  is number of output levels.

$$m_f = \frac{f_c}{f_m} \tag{4}$$

where:  $f_m$  is the modulation signal frequency,  $f_c$  is the carrier signal frequency.

In the case of the multilevel PWM each voltage level has its own carrier signal. The creation of gating signals for 7 level inverter as well as the inverter output voltage is shown in Fig. 3.

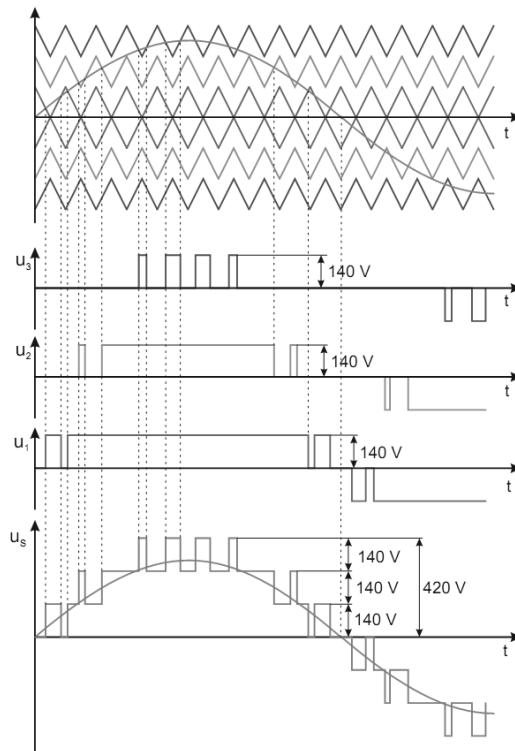


Fig. 3 Multicarrier opposition disposition PWM modulation for 7-level inverter

When increasing the number of output voltage levels based on (2) the number of carriers must be increased as well even though the number of H-bridges remains the same. The case for 15-level inverter is shown in Fig. 4. The control signals for virtual cascade inverter consisting of 7 H-bridge inverters are created in the same way as for the 7-level case in Fig. 3.

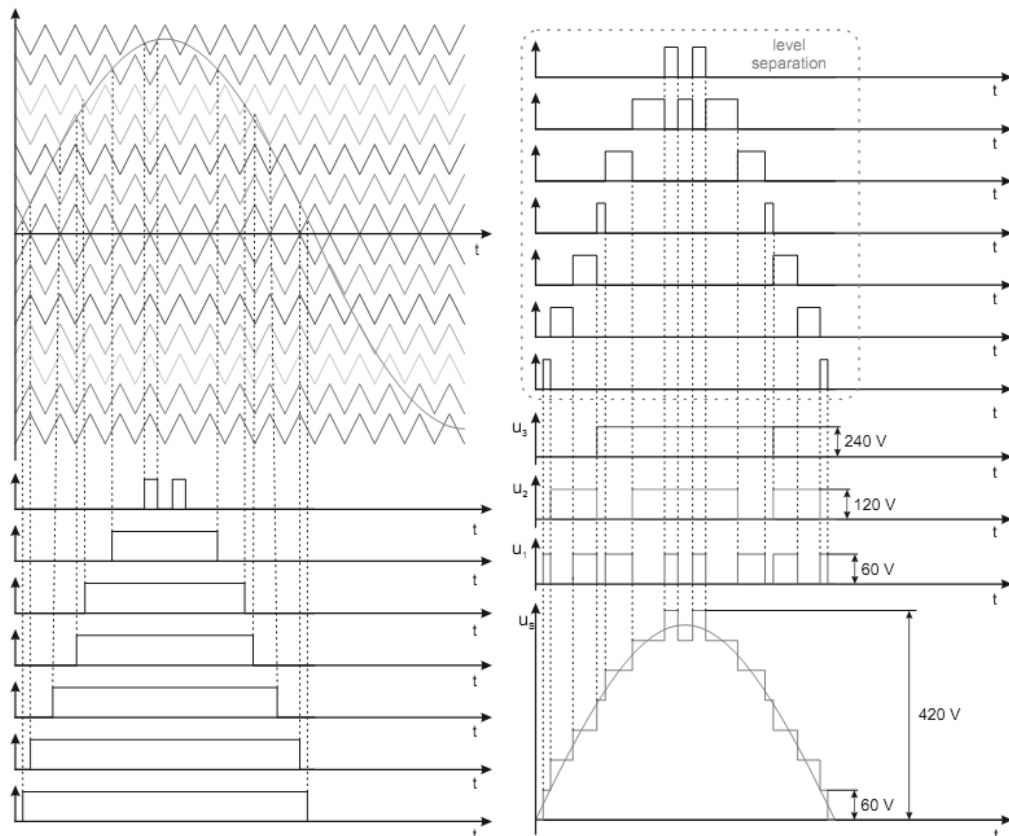


Fig. 4 Multicarrier opposition disposition PWM modulation for 15-level inverter

Because the real cascade inverter has only 3 H-bridge inverters it is necessary to convert the gating signals for 7 H-bridge inverters into the form suitable for 3 H-bridge inverters while the number of output voltage levels remains the same. This is done by level separation shown in Fig. 4. Level separation is done by logical function Exclusive OR which is applied to two neighboring voltage levels. The resulting three gating signals are used to control the laboratory model of the 15-level cascade inverter shown in Fig. 1.

### III. IMPLEMENTATION IN DSP

The multilevel PWM technique described in the previous section was implemented in DSP TMS320F28335 working with clock frequency of 150 MHz. The multilevel PWM is generated in the interrupt with time period of 10  $\mu$ s. The switching frequency of the cascade inverter was 5 kHz. One period of a carrier signal thus consists of only 20 steps (Fig. 5). This is rather low and the PWM thus has very poor resolution.

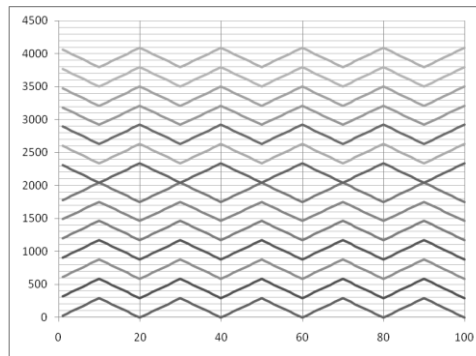


Fig. 5 Multilevel carriers for 15 level cascade inverter created in DPS

The sinusoidal modulation signal was created in RT-Lab as analog signal and was processed by ADC of the DSP. The connection to the laboratory model of cascade inverter was done through level shifter from 3,3 V to 5 V.

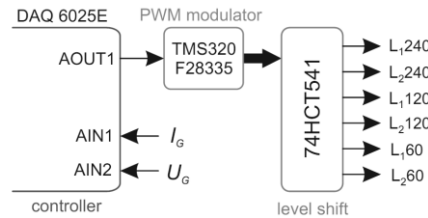


Fig. 6 Multilevel carriers for 15 level cascade inverter created in DPS

The example of the cascade inverter output voltage together with the partial H-bridges output voltages is shown in Fig. 7.

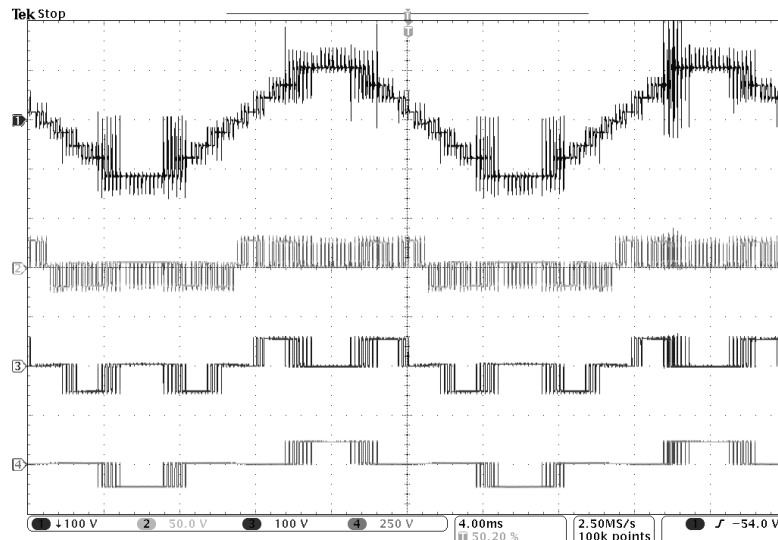


Fig. 7 Example of the multilevel inverter output voltage and H-bridge inverters output voltages

#### IV. CONCLUSION

A development process of the multicarrier sinusoidal PWM for multilevel inverter is presented. The PWM modulator was implemented in DSP. The algorithm is rather simple but takes many clock cycles. The DSP was able to create only low resolution PWM. Its limitations are not so visible in the voltage control. Low resolution PWM can be a problem when using the current control to control the inverter output current. This problem can be solved by using faster platform to perform simple logical operation, such as FPGA.

#### ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0185-10.

#### REFERENCES

- [1] L. M. Tolbert, "Novel Multilevel Inverter Carrier-Based PWM Method", *IEEE Transactions on Industry Applications*, Vol. 35, No. 5, p. 1098-1106, 1999
- [2] S. Laali, K. Abbaszadeh, H. Lesani, "Development of multi-carrier PWM technique for multilevel inverters," *Electrical Machines and Systems (ICEMS)*, pp.77,81, 2010
- [3] S. Sedghi, A. Dastfan, A. Ahmadyfard, "A new multilevel carrier based pulse width modulation method for modular multilevel inverter," *Power Electronics and ECCE Asia (ICPE & ECCE)*, pp.1432,1439, 2011
- [4] S. Manasa, R.V. Parimala, V. Chayapathy, "Advanced pulse width modulation techniques for cascaded multilevel inverters," *Sustainable Energy and Intelligent Systems (SEISCON 2011)*, pp.259,265, 2011
- [5] L.M. Tolbert, F.Z. Peng, T.G. Habetler, "Multilevel PWM methods at low modulation indices," *Applied Power Electronics Conference and Exposition, 1999. APEC '99, vol.2*, pp.1032,1038 vol.2, 1999
- [6] P. Palanivel, S.S. Dash, "Multicarrier pulse width modulation methods based three phase cascaded multilevel inverter including over modulation and low modulation indices", *TENCON 2009*, pp.1,6, 2009
- [7] S. Khomfoi, L.M. Tolbert, "Multilevel Power Converters", H. Rashid: *Power Electronics Handbook*, second edition, ISBN 0-12-088479-8, p. 451-482, 2007
- [8] S. Kouro, J. Rodriguez, J. Point, M. Angulo, "Multilevel Inverter Modulation method With DC-Link Disturbance Compensation", *IECON 2005*, ISBN 0-7803-9252-3, s.6, 2005

# Automatická segmentácia reči v programe Praat

<sup>1</sup>Matúš PLEVA, <sup>1</sup>Jozef JUHÁR, <sup>2</sup>Andrew Simon THIESSEN

<sup>1</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická Univerzita v Košiciach, Slovenská republika

<sup>2</sup>Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>Matus.Pleva@tuke.sk, Jozef.Juhar@tuke.sk, <sup>2</sup>thiessenandrew@gmail.com

**Abstrakt** – Program Praat je všestranný program na analýzu reči. Pre účely segmentácie reči bol pre Praat vyvinutý a použitý plugin s názvom EasyAlign. V tomto článku bude predstavených niekoľko rôznych nových funkcií implementovaných do programu Praat pre zlepšenie procesu segmentácie reči. Podobne bude opísaná integrácia systému na automatické rozpoznávanie reči a predspracovanie zvuku do programu Praat, kvôli rozšíreniu a zlepšeniu procesu segmentácie reči a využiteľnosti celého nástroja. Grafický výstup Praatu tiež pomáha vizualizovať výsledky segmentácie reči a automatického rozpoznávania spojitých reči na vzdelávacie účely.

**Kľúčové slová** – Praat, EasyAlign, segmentácia reči, rozpoznávanie reči

## I. ÚVOD

Segmentácia rečových nahrávok pred ich spracovaním je dôležitým problémom, ktorý je nutné riešiť pri spracovaní dlhých nahrávok spojitých reči. Segmentácia umožňuje spracovanie nahrávok či diktovaných úsekov dlhších ako 3 minúty, čo je väčšinou maximálna hranica pre spracovanie spojitého segmentu bez výraznejších páuz (tichých segmentov). Vďaka segmentácii môže samotný proces rozpoznávania reči byť paralelizovaný, kde každý segment spracuje iná inštancia modulu automatického rozpoznávania reči (ASR). Program Praat má veľký potenciál v oblasti segmentácie reči, vďaka mnohým nástrojom na spracovanie a analýzu reči s rôznym využitím a možnosťou skriptového programovania.

Táto práca skúma a opisuje implementáciu nových postupov segmentácie reči v Praate a rozširuje tiež existujúce funkcie. Nové funkcie Praatu<sup>1</sup> boli implementované použitím Praat skriptov, vďaka ktorým je možné kontrolovať, vylepšiť a rozšíriť možnosti spracovania reči. Tieto funkcie obsahujú hlavne rozšírenú segmentáciu reči (rozšírenie pluginu EasyAlign<sup>2</sup>), doplnenie rozpoznávania reči do takto segmentovaných úsekov, automatická segmentácia použitím nástroja *distbic* a následné využitie modulu ASR a nakoniec predspracovanie rečového signálu pomocou externého nástroja SoX<sup>3</sup> pre vylepšenie ďalšieho spracovania.

Téma segmentácie reči je esenciálna pre zlepšenie výsledkov automatického rozpoznávania reči a preto je dôležitá v dobe, kedy sa rečová komunikácia s počítačmi stáva viac a viac integrovaná do každodenného používania počítača, tabletov a chytrých telefónov.

Praat skripty boli naprogramované pomocou Praat skriptovacieho jazyka. Všetky softvérové pluginy a nástroje používané v tejto práci sú open source, a môžu byť voľne získané on-line, okrem nástroja *distbic*, ktorý vznikol našou implementáciou algoritmu [1] a novo vyvinutého servera rozpoznávania reči slovenského jazyka opísaného v [2].

## II. CIELE

Medzi hlavné ciele môžeme zaradiť: Po prvé, analyzovať súčasný stav akustickej segmentácie reči a funkcií použitých v EasyAlign pluginu. Analýza rôznych možností programovania v programe Praat a doprogramovania nových funkcií, ako je zníženie hluku/šumu v nahrávkach alebo expanzia dynamického rozsahu nahrávky, pomocou externých nástrojov. Takisto je v článku

<sup>1</sup><http://www.praat.org>

<sup>2</sup><http://latlcul.unige.ch/phonetique/easyalign.php>

<sup>3</sup><http://sox.sourceforge.net>



opísaná práca s objektami a skriptami v Praate a možnosti pridania nového TextGrid-u na vizualizovanie výsledkov z externých modulov (ako ASR, SoX, distbic). Ďalej, bude opísané vytvorenie novej verzie aplikácie segmentácie reči: pre segmentáciu reči ak existuje vopred textový prepis nahrávky (to čo bolo presne prečítané rečníkom), a verzia ktorá nemá žiadne dodatočné vstupné informácie o spracovávanej nahrávke a je plne nezávislá od EasyAlign pluginu. Bude takisto opísaná možnosť porovnania s výsledkami serveru pre automatické rozpoznávanie spojitkej reči v slovenčine, v prípade, že bude vstupný text dostupný.

### III. ANALÝZA SÚČASNÉHO STAVU

Segmentácia reči (alebo zarovnanie reči), je delenie a segmentácie reči vyslovene aj písanej do menších lexikálnych a fonetických častí. To sa vykonáva delením vety do slov, slov na slabiky a slabík na fonémy. Fonetická segmentácia (alebo fonetické zarovnanie) určuje pozíciu a poradie rôznych foném a slabík v reči. Zvuková audio nahrávka a textový (ortografický) prepis danej reči sa používa pri zarovnaní, aby sa pokúsil zladit' text a zvuk dohromady [3].

Ak chcete zvýšiť efektivitu Praatu (s EasyAlign pluginom) v jeho procese segmentácie reči, musíme uvažovať ďalšie aspekty reči. Základným spôsobom ako zlepšiť segmentáciu reči v Praat je využiť automatické rozpoznávanie reči, zvýšenie kvality zvuku a použitie externých nástrojov na segmentáciu reči v Praate.

Kvalita vstupného audia je veľmi dôležitá pre segmentáciu reči, pretože jej základná verzia analyzuje energiu signálu v rôznych frekvenčných oblastiach. Nahrávky so šumom v pozadí predstavujú veľkú prekážku pre správne určenie hranice slov podľa dodaného textu pri použití základnej segmentácie pluginu EasyAlign. Nie je možné zaručiť kvalitné nahrávky pri rôznych nahrávacích podmienkach či dokonca mobilných zariadeniach. Preto je vhodné požiť predspracovanie nahrávky aplikovaním šumového filtra alebo dynamickej expanzie signálu (expander), ktorá stíši tichšie segmenty a tak budú pre EasyAlign ľahšie identifikovateľné medzi slovami. Takéto spracovanie audia je preto vhodné pridať do Praatu ako voliteľnú možnosť pred použitím algoritmov segmentácie rečových nahrávok.

Použitie vzdialeného servera automatického rozpoznávania spojitkej reči (ASR/LVCSR) umožňuje porovnanie výsledkov tohoto servera s pôvodným textovým prepisom nahrávky a tým si overiť či ladiť funkčnosť servera v grafickom prostredí Praatu. Samozrejme je možné použiť ASR server aj bez vstupného textu, čo rozšíri možnosti segmentácie súborov, ktoré nemajú vopred známy rečníkom prečítaný text.

Pre zlepšenie akejkoľvek oblasti používania programu Praat musí byť použitý skriptovací jazyk Praat, či už vytvorením novej alebo upravením starých Praat skriptov. Praat spúšťa tieto skripty a oni riadia Praat a ďalšie externé programy, ktoré pracujú spoločne s ním. Skripty môžu byť vykonané z hlavného menu Praat priamo z rozhrania Praat-u [4] ak dôjde k dedefinovaniu konfiguračného súboru Praat menu.

Hlavným pluginom, ktorý bol použitý ako odrazový mostík pre ďalší vývoj v oblasti segmentácie reči v Praate bol plugin EasyAlign. Jedná sa o trojstupňový proces segmentácie reči, ktorý používa zvukový a textový súbor na vytvorenie Praat TextGrid-u výslednej segmentovanej reči. Všetky nové funkcie by mali byť úzko prepojené s EasyAlign pluginom v programe Praat [5].

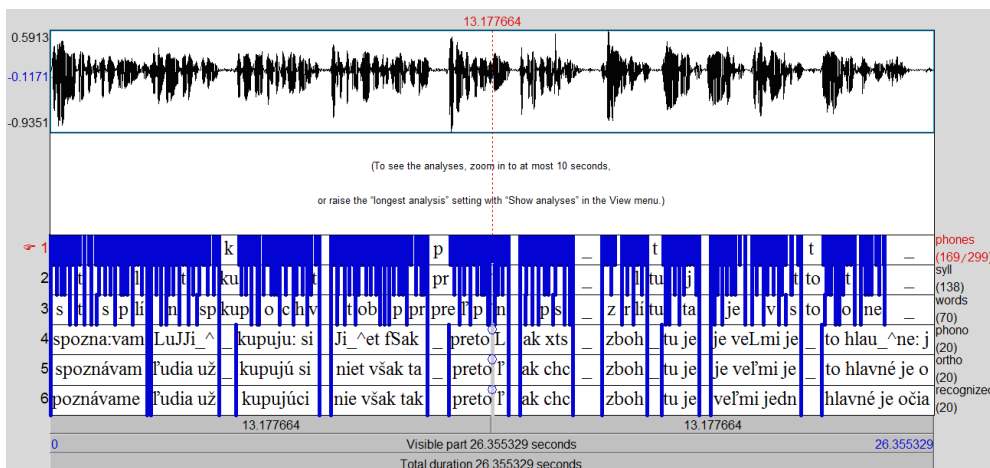
### IV. VÝSLEDKY

Implementácia nových funkcií s cieľom zlepšiť proces segmentácie reči v Praat-e bola dosiahnutá integrovaním niekoľkých externých podprogramov. Zároveň boli vytvorené štyri nové Praat funkcie. Patria medzi ne: nový krok rozpoznávania spojitkej reči s veľkým slovníkom (LVCSR) v pôvodnom EasyAlign procese segmentácie reči; nová funkcia LVCSR a segmentácie reči, ktorá dokáže spracovať zvuk bez vstupného textu; funkcie pre spracovanie viacerých zvukových súborov s použitím druhej novej funkcie; a funkcie predspracovania zvukového súboru (odšumenie a expander). Všetky novo implementované funkcie môžu byť spustené z menu EasyAlign v Praate.

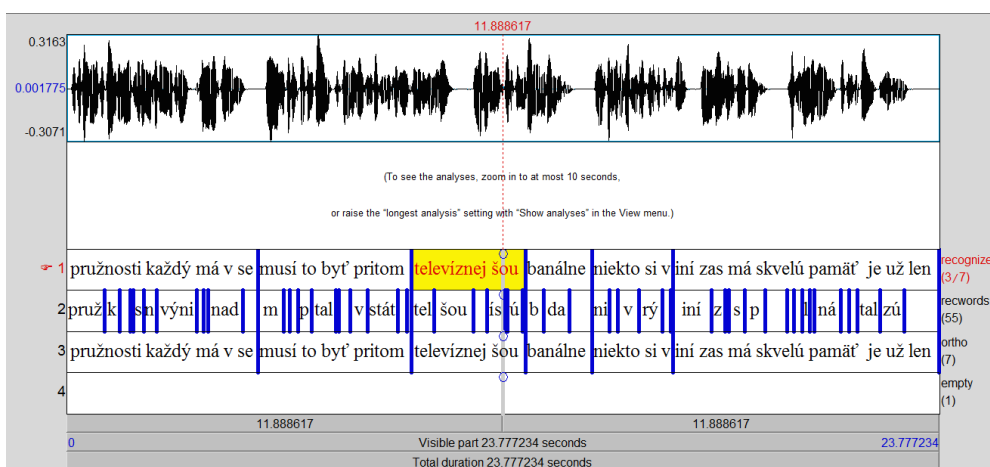
#### A. Implementácia rozpoznávania reči

Nový krok rozpoznávania reči bol implementovaný do EasyAlign pluginu dema pre segmentáciu reči. Tento krok umožňuje po skončení segmentačného dema spustiť automaticky proces ASR pre všetky vytvorené segmenty a zobrazit' výsledok vo forme jednoducho porovnateľnej s pôvodným vstupným textom. Každý interval - segment z nahrávky sa spracováva samostatne. Výsledkom je potom poslaný späť do Praat-u a vloží sa do nového *recognized* riadka.

Výsledkom je Praat-ovský TextGrid, ktorý zobrazuje segmentovaný a rozpoznávaný text, spolu so zvukovou stopou. Výsledný TextGrid je zobrazený na Obrázku 1.



Obr. 1 Praat TextGrid s výsledkami novej funkcie automatického rozpoznávania reči v *recognized* tier.



Obr. 2 Praat TextGrid výsledkov funkcie „rozpoznaj aj segmentuj“

### B. Implementácia funkcie „rozpoznaj aj segmentuj“

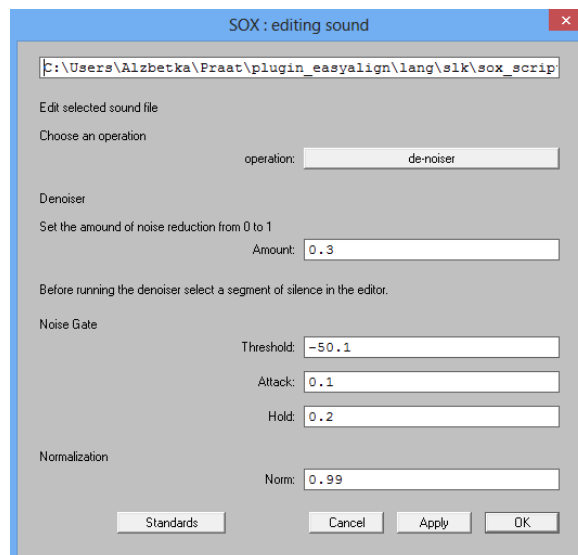
Funkcia „rozpoznaj aj segmentuj“ má tú výhodu, že nepotrebuje pre svoju funkčnosť segmentáciu EasyAlign pluginu, ktorý vyžadoval spolu s audiom aj vstupný text. Táto funkcia využíva vlastný segmentačný externý program *distbic*, ktorý vznikol už predtým implementáciou dvojfázového DistBic algoritmu opísaného v [1]. Takto vytvorené segmenty sú poslané do modulu ASR [6] a výsledok je po parsovaní vložený do TextGridu, spolu so slovným zarovnaním (ak bolo pri spustení funkcie zvolené). Ďalej je možné využiť funkcie EasyAlign pluginu na zobrazenie fonetickej transkripcie a fonetického a slabikového zarovnania. V module ASR sa používa jazykový model adaptovaný na spravodajské relácie [7] a tak je dostatočne všeobecný pre danú úlohu. Výsledok je zobrazený v TextGride (viď Obr. 2) ale takisto je uložený vo formáte *srt* titulkov, pre ďalšie spracovanie.

### C. Implementácia funkcií audio predspracovania

Funkcie audio predspracovania boli implementované aby umožnili užívateľovi vylepšovať kvalitu audia priamo v prostredí Praat-u. Nakoniec boli implementované tri audio nástroje pre predspracovanie: odšumenie, dynamická expanzia (gate) a normalizácia.

Funkcia odšumenia a dynamickej expanzie (gate) bola implementovaná do Praatu s využitím externého programu *Sound eXchange (SoX)*. Funkcia odšumenia eliminuje zvolené frekvenčné spektrum na základe užívateľom definovanej šumovej oblasti. Toto je aj jej hlavnou nevýhodou - potrebuje manuálne označenie šumového segmentu, a po aplikovaní odšumenia sú niekedy rečové zložky deformované nevhodným spôsobom pre ASR modul, avšak vhodnejšie pre segmentačný *distbic* modul.

Funkcia dynamickej expanzie (gate) poskytuje možnosť znížiť energie segmentov, ktoré identifikuje ako šumové vďaka prednastaveného threshold-u (hranici). Takto upravená nahrávka opäť pomôže hlavne segmentačnému *distbic* modulu získať presnejšie hranice segmentov.



Obr. 3 SoX menu

Implementovaná funkcia normalizácie poskytuje možnosť priamo v Praate zvýšiť hlasitosť nahrávky na maximum bez akéhokoľvek znehodnotenia nahrávky. Funkcia najprv analyzuje nahrávku a zvýši hlasitosť tak aby najhlasnejší úsek bol na maxime rozsahu (0.99) a o rovnaký koeficient zvýši hlasitosť celej nahrávky: *Scale peak*.

V Praat-e bolo vytvorené špeciálne SoX menu, kde môže byť zvolená funkcia aj s jej parametrami (viď Obr. 3) a po jej skončení sa vytvorí nový zvukový objekt so spracovaným zvukom.

## V. ZÁVER

Cieľom tejto práce bolo zlepšenie programu Praat pre použitie na segmentáciu reči, čo bolo dosiahnuté využitím množstva rôznych nových metód, postupov, externých programov a Praat funkcií. Tým, že kombinuje proces segmentácie reči, automatického rozpoznávania spojitých reči a predspracovania zvuku bolo dosiahnuté podstatné vylepšenie funkcionality programu Praat a segmentácie reči pluginom EasyAlign, ktorý je užitočný v procese slabikovej a fonetickej segmentácie a prehľadnej prezentácie výsledkov.

Takisto významným vylepšením je možnosť zobrazit' výsledky externého ASR modulu spolu s originálnym audiom, prípadne pôvodným textovým prepisom, čo umožňuje prehľadne, jednoducho a účinne testovať externý ASR systém.

Vďaka využitiu Praat skriptov je rozšírenie ľahko použiteľné v ďalšom výskume v danej oblasti.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Kompetenčné centrum znalostných technológií pre inovácie produkčných systémov v priemysle a službách (kód ITMS projektu: 26220220155) 50% a projektu Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií (ITMS-26220220182) 50% na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] J. Žibert et al, "COST278 broadcast news segmentation and speaker clustering evaluation." *Proc. of Interspeech*, 2005, Bonn, pp. 629-632.
- [2] M. Lojka, S. Ondáš, M. Pleva a J. Juhár, "Multi-thread parallel speech recognition for mobile applications." *Journal of Electrical and Electronics Engineering*, Vol. 7, no. 1 (2014), p. 81-86, eISSN: 2067-2128.
- [3] J.P. Goldman, EasyAlign: An Automatic Phonetic Alignment Tool Under Praat. *Proc. of InterSpeech*, 2011, Florence, p. 3233-3236.
- [4] W. Styler, *Using Praat for Linguistic Research*. University of Colorado at Boulder Phonetics Lab, Document version 1.4.2, 16 January 2014.
- [5] J.P. Goldman, *Tutorial on EasyAlign*. University of Geneva [Online], p. 1-4.
- [6] S. Ondáš, M. Pleva, M. Lojka, M. Sulír a J. Juhár, "Server-based speech technologies for mobile robotic applications." *Journal of Electrical and Electronics Engineering* [online]. May 2013, vol. 6, iss. 1, p. 95-98. ISSN 1844-6035.
- [7] M. Pleva a J. Juhár, "TUKE-BNews-SK: Slovak Broadcast News Corpus Construction and Evaluation." *Proc. of LREC*, May 2014, Reykjavik, Iceland, pp. 1709-1713.

# Impulzný ohrev, ako možnosť úpravy magnetických vlastností feromagnetických amorfných zliatin

<sup>1</sup>Lukáš HUBAČ, <sup>2</sup>Ladislav NOVÁK

<sup>1,2</sup>Katedra fyziky, Fakulta elektrotechniky a informatiky, Technická univerzita v  
Košiciach, Slovenská Republika

<sup>1</sup>lukas.hubac@tuke.sk, <sup>2</sup>ladislav.novak@tuke.sk

**Abstrakt** – Zdokonalenie magnetických vlastností amorfnej zliatiny typu FINE-MET v tvare pásky bolo skúmané pomocou novo navrhnutého zariadenia na kontrolovaný pulzný ohrev vzorky, ktoré umožňovalo súčasnú aplikáciu mechanického napätia. Žihací striedavý prúd mal hodnotu cca 21 A a frekvenciu 50 Hz. Každá vzorka bola ohrievaná sériami pulzov rôznej dĺžky od 80 ms do 120 ms priamo v solenoide meracej aparatury magnetometra, pomocou ktorého sa realizovali následné magnetometrické merania. Umiestnenie v solenoide umožnilo realizovať aj žihanie pod vplyvom vonkajšieho poľa. Teploty dosahované pri žihaní boli do 400 °C. Magnetické parametre ako koercivita a konštanta celkovej anizotropie boli skúmané v závislosti na počte impulzov. Zistili sme, že pri pulznom žihaní sériami pulzov (20 opakovaní v sérii) s danými parametrami sa nevytvorili nanometrické zrná ani koncentrované distribúcie, no magnetické vlastnosti sa výrazne zlepšili.

**Kľúčové slová** – FINEMET, impulzné žihanie, koercivita, anizotropia.

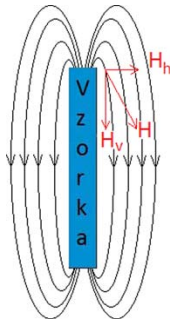
## I. ÚVOD

Nanokryštalickou transformáciou amorfného FINEMET-u vzniká materiál, ktorý má široký rozsah využitia ako magneticky mäkký materiál. Nevýhodou nanokryštalického materiálu je však jeho krehkosť. Z tohto dôvodu nie je vhodný pre aplikačné účely na miestach aj s malými otrasmami a vibráciami. Napríklad v automobilovom a leteckom priemysle. Preto je našim úmyslom vylepšovať magnetické vlastnosti pri zachovaní amorfného stavu. Tradičné tepelné spracovanie je vo veľkom meradle využívané v priemyselnej výrobe na nanokryštalizovanie tejto zliatiny [1], [2]. Pokúsili sme sa preskúmať vplyv impulzného žihania na magnetické vlastnosti tohto materiálu (koercivita a magnetická anizotropia) v takom rozsahu, aby nedošlo k vzniku nanokryštalickej štruktúry.

## II. POPIS MERACEJ APARATÚRY A APARATÚRY NA IMPULZNÝ OHREV

### A. Meracia aparatura na magnetometrické merania

Na meranie magnetickej polarizácie vzorky je použitá magnetometrická metóda. Pri tejto metóde je magnetická polarizácia vzorky stanovovaná prostredníctvom vyhodnocovania horizontálnej zložky rozptylových polí vzorky. Jedná sa teda o meranie magnetických vlastností otvorených vzoriek. Pri tomto druhu merania dostávame informáciu z celej dĺžky vzorky. Výhodou tejto metódy je, že umožňuje zmerať magnetickú polarizáciu vzorky bez zmeny magnetického stavu vzorky. Nevýhodou je, že pri otvorených vzorkách treba počítať s demagnetizačným faktorom, spôsobeným existenciou rozptylových polí. Zmagnetovaná otvorená vzorka vytvára okolo seba magnetické pole, ktoré sa nazýva rozptylové pole. Veľkosť tohto poľa je úmerná magnetickej polarizácii vzorky. V ľubovoľnom bode, okrem miesta, v ktorom sú magnetické indukčné čiary rovnobežné s osou vzorky, môžeme intenzitu rozptylového magnetického poľa vzorky rozložiť na jeho horizontálnu zložku  $H_h$  a vertikálnu zložku  $H_v$ , obrázok 1.

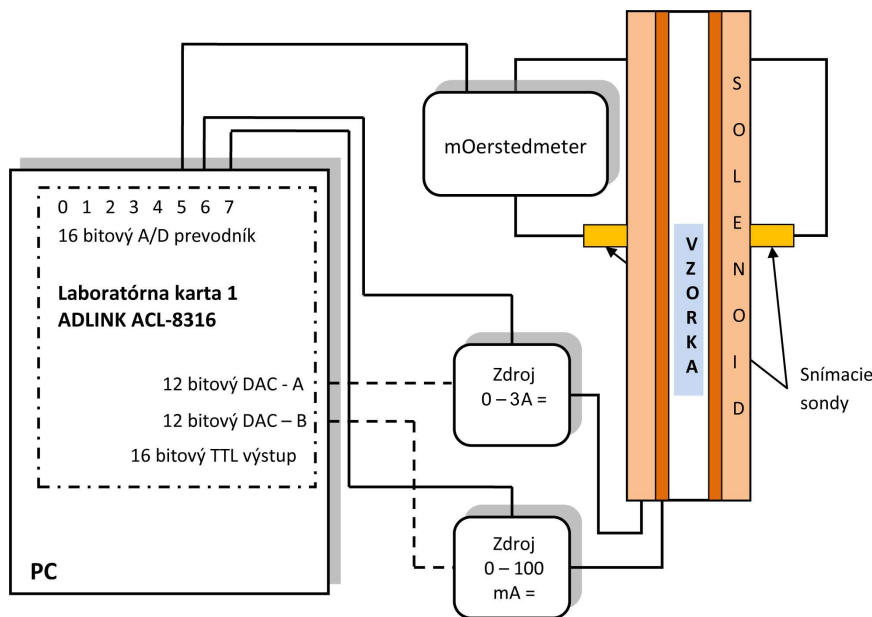


Obr. 1: Rozptylové magnetické pole vzorky

Takže aj intenzita horizontálnej zložky rozptylového magnetického poľa, meraná v konkrétnom bode, je úmerná magnetickej polarizácii vzorky. Konštantu úmernosti dostaneme, zmeraním signálu kalibračnej cievky, vloženaj na miesto vzorky. Na obrázku 2 je schematicky znázornený magnetometer, ktorý sa používa na Katedre fyziky TU v Košiciach. Rozptylové polia vzorky sú snímané dvoma súčtovými sondami komerečného milioerstedmetra. Snímacie sondy sú uložené oproti sebe v polovici výšky solenoidu. Vzorka je v solenoide umiestnená tak, aby jeden z jej pólov bol na úrovni sond. Dôvodom takého usporiadania experimentu je dosiahnutie čo najvyššej citlivosti zariadenia [3].

Uvedený spôsob merania magnetickej polarizácie je výhodný z niekoľkých dôvodov:

- Dvojica sond umiestnená proti sebe umožňuje precízne meranie s nízkou hladinou porúch.
- Je možné určiť magneticú polarizáciu vzorky absolútne (napr. po demagnetovaní vzorky) [3].
- Sondy sú umiestnené na vonkajšej strane solenoidu a celé vnútro solenoidu je voľné a možno v nom umiestniť napr. piecku na teplotné spracovanie a meranie teplotných závislostí, alebo zariadenie na impulzný ohrev a pod.



Obr. 2: Bloková schéma magnetometra [4]

### B. Popis aparatury na impulzný ohrev a určenie teploty dosahovanej počas impulzu

Existujúce zariadenie na meranie magneticých vlastností bolo rozšírené o zariadenie na realizáciu impulzného ohrevu vzorky priamo v meracom zariadení.

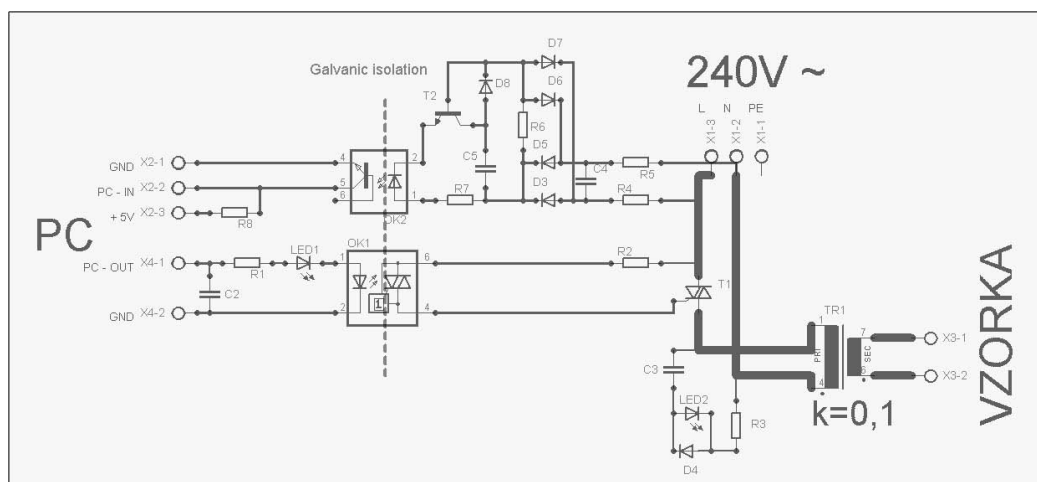
Táto metóda má niektoré výhody pre vyšetrovanie výsledkov impulzného žihania. Napríklad, dĺžku impulzu možno ovládať v širokom časovom rozsahu od 0,02 s až po celé sekundy. Čo umožňuje dobre regulovať teplota.

Vzorka sa vkladá vertikálne do vertikálne orientovaného solenoidu s vnútorným priemerom 40 mm a výškou 500 mm z vrchu v špeciálnom držiaku. Tento držiak je usposobený tak, aby vzorka mohla byť žihaná a zároveň vertikálne zaťažovaná rôznymi silami v tahu. Vzorka môže byť žihaná aj v magneticom poli.

Zariadenie na impulzný ohrev sa skladá zo špeciálneho držiaka, výkonovej a riadiacej časti. Táto sa skladá z nasledujúcich častí:

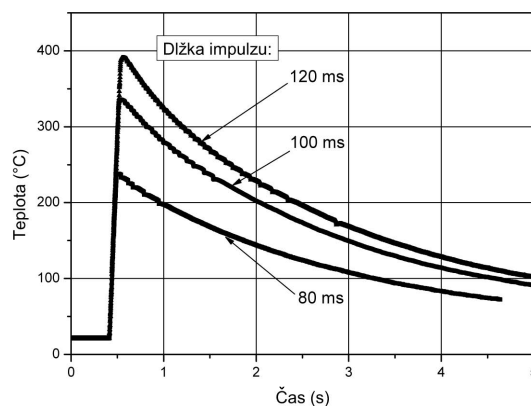
- počítačom riadený impulzný zdroj,
- snímanie počtu pulzov,
- riadiaci program.

Impulzný zdroj sa skladá z riadiacej a výkonovej časti, ktoré sú od seba galvanicky oddelené optočlenom (opto-diak). Výkonová časť je tvorená výkonovým triakom a transformátorom, ktorého transformačný pomer je cca 0,1 s výkonom 250 W. Primárne vinutie transformátora je riadené týmto triakom, ktorý je spínaný opto-diakom (MOC 3041) v nule. Optický člen je ovládaný počítačom cez laboratórnu kartu ADLINK ACL-8316 prostredníctvom TTL logiky. Na sekundárne vinutie transformátora je priamo pripojená vzorka. Prechodom prúdu cez vzorku (cca. 21 A) dochádza k joulovmu ohrevu. Na to, aby bolo možné riadiť počet impulzov, bolo potrebné rozšíriť zariadenie o spätnú väzbu. Výstupom tejto spätnej väzby v podobe TTL logiky je detekovanie prechodu striedavého prúdu nulou. Tento signál je privádzaný do riadiaceho počítača, kde sa prostredníctvom programu riadi celý joulov ohrev a následné meranie magnetických vlastností vzorky. Schéma zapojenia impulzného zdroja a detektora prechodu striedavého prúdu nulou je na nasledujúcom obrázku 3 [4].



Obr. 3: Schéma zapojenia impulzného zdroja

Ďalšou úlohou bolo určiť teploty dosahované pri jednotlivých impulzoch. Pôvodne sme chceli toto meranie realizovať umiestnením IR (Infra Red) snímačej diódy do solenoidu k vzorke, aby sa dala teplota merať priamo počas pulzu. Keďže každá polovodičová IR-dióda, ktorú sme mali k dispozícii, bola vyrobená z magnetického materiálu, tak sme pristúpili k okalibrovaniu jednotlivých impulzov. To bolo realizované bezkontaktné IR miniatúrnym pyrometrom (CTF-SF15-C3 MICRO-EPSILON). Výhodou bezkontaktného merania teploty pomocou daného pyrometra, boli veľké tepelné kapacity a malé reakčné časy snímacích čidiel. Rýchlosť tepelného ohrevu bola cca 3000 °C/s. Ochladzovanie prebiehalo po exponenciálnej krivke rýchlosťou v priemere 100 °C/s. Priebehy jednotlivých impulzov sú na obrázku 4 [5], [6].



Obr. 4: Grafická závislosť teploty od času pri jednotlivých dĺžkach impulzov. Impulz začína v čase 0,5 s.



Magnetické parametre: koercivita  $H_c$  a konštanta celkovej magnetickej anizotropie  $K_i$ , boli stanovené z následne meraných hysteréznych slučiek a to v pôvodnom stave vzorky a po každom impulze. Počet impulzov bol 1 až 20.

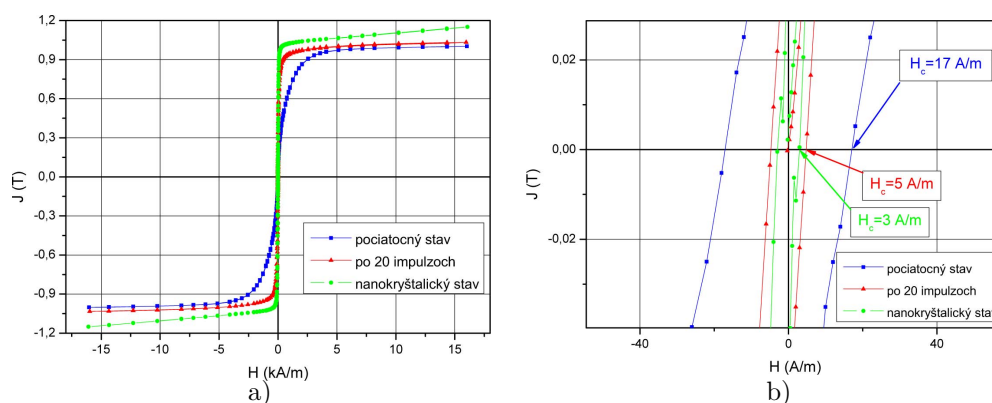
### III. EXPERIMENTÁLNE VÝSLEDKY A DISKUSIA

Pre štúdium uvedenej problematiky bol vybraný materiál známy ako FINEMET ( $Fe_{73,5}Si_{13,5}B_9Nb_3Cu_1$ ), pripravený v KFKI MTA v Budapešti procesom rýchleho chladenia taveniny. Jedná sa o amorfný materiál v podobe pásky hrúbky  $30 \mu m$  a šírky  $10 mm$ . Magnetické merania boli uskutočnené v magnetometri fungujúcom na princípe merania rozptylových polí vzorky (II.A). Z nich boli stanovené magnetické parametre ako koercivita a konštanta celkovej magnetickej anizotropie.

Koercivita  $H_c$  bola určená ako priesečník hysteréznej slučky s horizontálnou osou poľa  $H$ . Konštanta celkovej magnetickej anizotropie  $K_i$  bola získaná ako plocha nad krivkou prvotnej magnetizácie, pomocou (1).

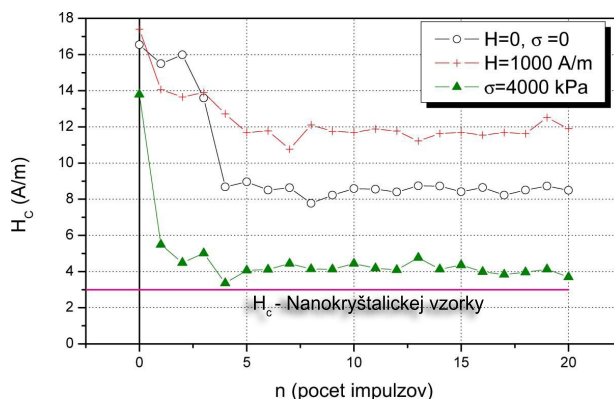
$$K_i = \int_0^{J_s} H \cdot dJ. \quad (1)$$

Na obrázku 5a môžeme vidieť, ako sa výrazne zmenila hysterézná slučka po 20 impulzoch s dĺžkou impulzu 120 ms. Došlo k zmenšeniu plochy slučky, a tým k zníženiu strát v procese premagnetovávania. Je zrejmé, že so zmenou strmosti hysteréznej slučky sa mení aj konštanta celkovej magnetickej anizotropie  $K_i$ . Obrázok 5b detailnejšie zobrazuje pokles koercivity po 20 impulzoch o viac ako dve tretiny.



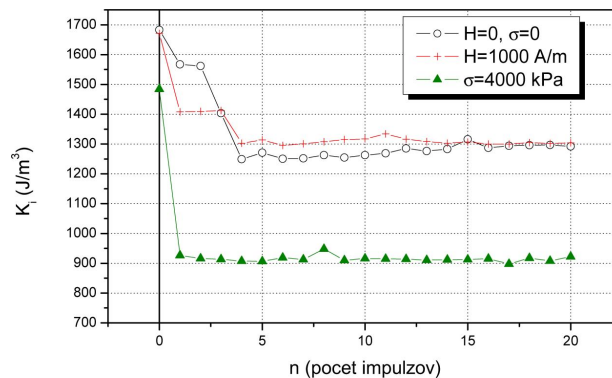
Obr. 5: Hysterézná slučka vzorky v počiatocnom stave a po sérii 20-tich impulzoch. Dĺžka jedného impulzu bola 120 ms. Vpravo je zobrazený detail pre zviditeľnenie poklesu  $H_c$ .

Obrázok 6 zobrazuje priebehy koercivity  $H_c$  v závislosti na počte impulzov a pri mechanickom namáhaní ( $\sigma = 0$  a  $4000$  kPa). Po sérii impulzných ohrevov dochádza k výraznému poklesu koercivity  $H_c$ . Toto možno vysvetliť v prvom rade výrazným znížením vnútorných napätí vnesených do materiálu v procese výroby.



Obr. 6: Zmena koercivity  $H_c$  v závislosti na počte impulzov a pri mechanickom namáhaní. Kde  $H$  je vonkajšie magnetické pole a  $\sigma$  je jednoosové axiálne mechanické namáhanie.

Konštanta celkovej magnetickej anizotropie  $K_1$  sa po sérii impulzov takisto výrazne zmenší v dôsledku poklesu vnútorných pnutí, ako je znázornené na obrázku 7.



Obr. 7: Zmena celkovej magnetickej anizotropie  $K_1$  v závislosti na počte impulzov a pri mechanickom namáhaní.

#### IV. ZÁVERY

- 1) Bol navrhnutý a skonštruovaný zdroj na impulzný ohrev.
- 2) Jednotlivým impulzom ohrevu bola meraním priradená dosahovaná teplota (okaličovanie impulzu).
- 3) Impulzným ohrevom bolo dosiahnuté výrazné zlepšenie magnetických vlastností materiálu (cca o 70%), bez vytvorenia nanokryštalickej štruktúry a vzorka ostáva mechanicky pružná aj po impulznom spracovaní (vzorka sa neláme).

Čo možno vysvetliť nasledovne:

- a) V procese impulzného ohrevu dochádza k difúzii atómov na krátku vzdialenosť, čo stačí na výrazné zníženie vnútorných napätí.
- b) Krátka doba impulzu však nedovolí, aby prebehla difúzia na dlhú vzdialenosť, ktorá by mohla viesť k vytváraniu klastrov, prípadne kryštálov.

#### POĎAKOVANIE

Tento článok vznikol za podpory vedeckej agentúry VEGA v rámci riešenia projektu VEGA 1/0778/12.

#### LITERATÚRA

- [1] A. Lovas and et al., *J. Phys. IV (France) 8.*, vol. Pr2/298, 1998.
- [2] L. Varga and et al., *J. Magn. Magn. Mater.*, pp. 477–479, 2003.
- [3] L. Novák, *Moderné magnetické mäkké materiály*. Košice, Slovakia: TUKE, 2007.
- [4] L. Hubáč, “Device for pulse annealing of amorphous magnetic ribbon,” *SCYR 2013*, pp. 250–252, May 2013.
- [5] L. Hubáč, “Pulse annealing of FINEMET type amorphous ribbons,” *SCYR 2014*, pp. 50–51, May 2014.
- [6] L. Hubáč and L. Novák, “Magnetické vlastnosti amorfnej zliatiny typu FINEMET v procese spracovania impulzným ohrevom,” *Senzorika a magnetometria 2013*, pp. 26–30, December 2013.



# Methods of student knowledge visualisation

<sup>1</sup>Peter SEMAN, <sup>2</sup>František JAKAB

<sup>1,2</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>peter.seman.2@student.tuke.sk, <sup>2</sup>frantisek.jakab@tuke.sk

**Abstract** — The main goal of this work is to create the application which will be able to get contents of tests from remote systems and visualize them in on-line graphs and diagrams. In analysis part we take a closer look at LMS Moodle where we point on its properties and absences. After that are introduced progressive forms of visualization techniques, their actual standing and importance of their existence. In the next chapter is designed a concept of future application and API (application programming interface) for targeted system. There are also mentioned used resources and extensions for its development and the creation of the architecture according to concept and modeling the whole application with all required additions.

**Keywords** — Diagrams, Evaluation, Graphs, Learning Management systems

## I. INTRODUCTION

Almost every school and learning institution uses the web based application to manage school content. They used to be called Learning management systems (LMS). There is in those applications everything under one roof. The function of the roof is to create a course. Which usually consists of on-line materials such as tests, assignments, evaluations, gradebooks and many other things directly connected to the education.

Few of them cover almost all the requirements of schools (or at least trying to), but many of them still have missing properties which lead to worse educational and/or evaluating processes for teachers.

They don't have implemented application programming interface (API) which would allow setup the remote connection. This functionality might be very useful, because many third service applications would be able to communicate with desired system and fetch/put the appropriate data.

Usually the missing part for an evaluation is visualization of results. This is also the main goal of this work. To create the application which can handle following actions to achieve fast and safe connection in order to get the proper results to visualize.

## II. MISSING FUNCTIONALITY OF NOWADAYS SYSTEMS

When we take a look on computerized testing and its evaluation, we can see the progress in consideration of previous years. Students do the tests on-line in few minutes intervals and teachers are sometimes not even sure why the results are like they are. Today exists many ways to evaluate a student's tests and their grading. We would like to create the better one - using visualisation graphs.

### A. CMS, LMS, Moodle

The Internet brings many on-line applications, which can be used for free. Users who use them are fully involved in their management processes and have bunch of permissions to achieve certain points. Mentioned types of applications are CMS which stands for Content Management Systems. They store information about published content and stuff that users created. [1] Also, there is very similar type of on-line application which is called Learning management system (LMS) and its goal is to let users manage their classes on-line. Most famous LMS is Moodle.

Despite the fact the science is moving forward the article about Moodle describes following

lacks:

1. Option to integrate dynamic visualisations a simulations of algorithms
2. Let the system automatically check the accuracy of submitted assignment
3. Availability to create student groups
4. Discussion right in studying material [2]

There are many researches which try to find new methods to accomplish an ideal combination of classic and on-line approach. They want to achieve "Hybrid course". It is virtual connection of computer and real world, where many of common things needed to learn are not required [3].

### B. Visualisation and its techniques

History of visualisations is based on simple drawings in caves, but today it represents expansive applications in various disciplines: computer engineering, education, science and also in medicine. It's important to know that under 'visualisation' word we shouldn't stay in static context, but it has been spread also in dynamic context - to the animations [4].

The most used diagrams in today's applications are following: Column Chart, Bar Chart and Pie Chart. They are very easy to implement and can absorb pile of data which is quite simple visualised.

But Gantt Charts, TreeMapBar Chart, Filled Donut Chart are advanced visualisation charts, which can absorb doubled or tripled size based on basic charts. The reason is that they can fill every free space in their content, so the purpose of visualisation is really maximized. For further imagination, please note Fig. 1 below, which represents TreeMapBar Chart which is filled to the last available pixel.

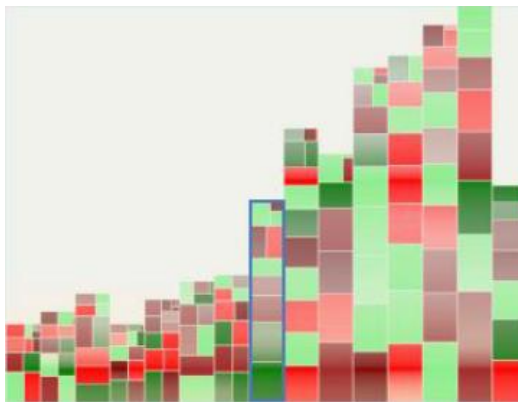


Fig. 1 Example of TreeMapbar Chart

### III. SUGGESTED SOLUTION AND RESULTS

Based on known information we created an application which should take care about missing parts from related systems.

Application concept is based on API calls. There will be existed remote system (Moodle) and internal system, which would handle communication with API. We can design concept like it is visible on Fig. 2.

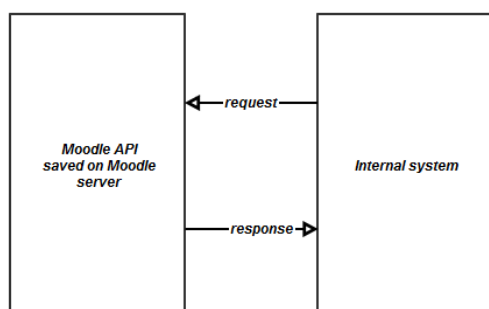


Fig. 2 Architecture design

Created application is defined as the system consisting of simple interface which allows logged users to maintain the remote content from connected API. Basically logged user has to create at least one system, which is represented by:

- API url for Moodle API
- API url and Token for Canvas REST API

In case of choosing Moodle API, one-time authentication and authorization is needed in running session. Establishing of connection is not required while Canvas REST API is chosen, because token did the work. Interface of application is quite simple and logged user is immediately ready to generate the desired content.

So after implementing the functionality there is little updated architecture considering its design. Existing architecture (Fig. 3.) looks as follows:

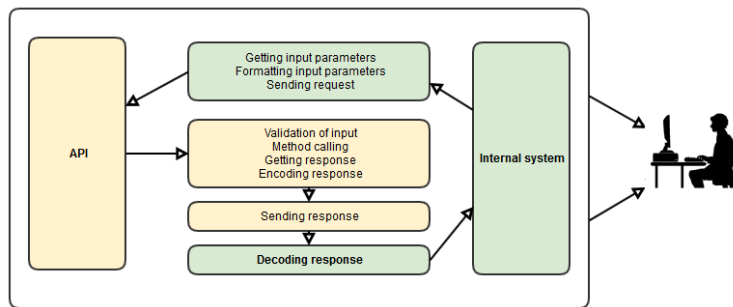


Fig. 3 Architecture (bridge between internal application and API)

Flow of generating diagrams (Fig. 4.) can be written into these steps:

1. Choice of course
2. Choice of test
3. Choice of date ranges
4. Choice of type of diagram and type of report
5. Generated diagram is clickable, so it's possible to click on some part of it
6. Choice of detailed report for selected part of diagram
7. Sometimes detailed generated diagram is clickable, so it's possible to click on some part of it

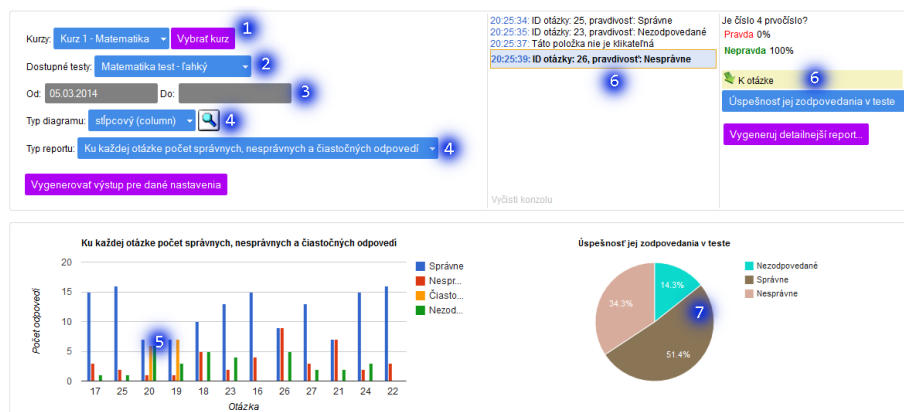


Fig. 4 Flow of generating diagrams

As Fig. 4. says - points 5,7 - there is available traversal through generated graphs. It is functionality which covers the main idea of this application – so user is able to see different views (graphs) very fastly. User can click to every place and the console is the place giving the output for clicks. So users are allowed to manage generating flow on their own, because the

responsibility of generated content depends on them. Also, when graph is generated, once the page is changed/switched, graph is gone. To prevent "losing" data there is feature to export reports to portable formats. These formats are:

1. Pdf report - covers all the data from the screen (course, test, filters, diagram type, graph, description)
2. Png image
3. DataTable - this format was created in order to using graph data again.

All of mentioned reports above are saved on the server side, so teacher might come back to appropriate menu "My reports" and there are visible his reports divided into three known categories available to immediate download.

#### ACKNOWLEDGMENT

This Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] CMS (Content Management System) Definition 29.10.2013, available on-line <http://www.techterms.com/definition/cms> G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
- [2] G. Rößling, P. Crescenzi, P. Ihantola, M. McNally, A. Radenski, M. Gloria Sánchez-Torrubia, *Adapting Moodle to Better Support CS Education*, 2010.
- [3] J. Cole, H. Foster, *Using Moodle, 2<sup>nd</sup> Edition, Second Edition*. O'Reilly, 2007, pp 18-19, ISBN 0-596-52918-X
- [4] V. Voželínek, *Cartography for GIS – Geovisualization Map Communication*, 2005 Olomouc, ISBN 80-244-1047-8

# Klasifikácia názorov dolovaných v konverzačnom obsahu

<sup>1</sup>Martin MIKULA <sup>2</sup>Kristína MACHOVÁ,

<sup>2</sup>Katedra kybernetiky a umelej inteligencie, FEI, Technická Univerzita Košice, Letná 9, 042 00, Košice

<sup>1</sup>Martin.Mikula@tuke.sk, <sup>2</sup>Kristina.Machova@tuke.sk

**Abstrakt** – Tento článok je venovaný klasifikácii názorov dolovaných v konverzačnom obsahu. V prvej časti sa venuje vysvetleniu, pojmu konverzačný obsah, vysvetleniu čo je to klasifikácia názorov. Ďalej sa venuje postupu pri zisťovaní názorovej polarity a predspracovaniu textu, ktorý sa bude analyzovať. Teoretická časť je venovaná spôsobom, akými sa v súčasnosti prístupuje k spracovaniu názorovej polarity a ako sa postupuje pri tvorbe slovníkov. V praktickej časti práce je popísaný návrh a implementácia nového slovníkového prístupu k identifikácii názorovej polarity webovej diskusie. Pre efektívnu prácu algoritmu boli vytvorené aj 4 rôzne slovníky, ktoré boli rozlíšené na základe typov slov, ktoré sa v nich nachádzali. Na konci sú popísané jednotlivé testy, ich vyhodnotenie.

**Kľúčové slová** – Klasifikácia názorov, webová diskusia, slovníkový prístup

## I. ÚVOD

Súčasný sociálny web umožňuje interakcie medzi veľkým množstvom ľudí, ktorí navzájom komunikujú pomocou komunikačných kanálov, ako sú napríklad chaty, blogy a mikro-blogy, diskusné fóra a rôznych ďalších. Tieto interakcie vplývajú nie len na ľudí, ktorí ich vykonávajú, ale aj na tých, ktorí ich čítajú. Tým majú možnosť meniť ich rozhodnutia v reálnom živote. Niekedy sú však tieto diskusie príliš dlhé a ich časti nie sú až také dôležité. Preto sú čím ďalej, tým viac populárne nástroje na analýzu takýchto neštruktúrovaných textov.

Klasifikácia názorov (sentiment classification) je proces, pri ktorom sa zisťujú názory, postoje a emócie používateľa k danej téme. Postoj je určený hodnotiacim faktorom (môže byť pozitívny alebo negatívny) a silou. Sila hodnotiaceho faktora je závislá od stupňa intenzity slov, viet alebo fráz, ktoré sa týkajú danej témy a ich počtu. Pod pojmom téma môžeme rozumieť akékoľvek hodnotenie produktov, osôb, filmov, kníh a podobne. Autorom respektíve držiteľom názoru je podľa [1] osoba, popríklad organizácia, ktorá má konkrétny názor na konkrétny objekt. Objekt je definovaný ako téma, na ktorú sa daný názor vzťahuje. Názor je teda postoj, pohľad alebo hodnotenie objektu autorom. Klasifikácia názorov je súčasťou dolovania názorov (opinion mining). V rámci dolovania názorov môžeme názory okrem klasifikácie aj vyhľadávať alebo inak analyzovať. Spolu s dolovaním názorov sa v literatúre objavuje pojem analýza sentimentu (sentiment analysis) [2]. Analýza sentimentu je však všeobecnejší pojem ako dolovanie názoru, lebo v rámci sentimentu môžeme analyzovať okrem názoru aj emócie.

## II. KLASIFIKÁCIA NÁZOROV

Pre klasifikáciu názorov nie sú potrebné všetky slová. Zaujímavé sú iba slová, ktoré sú nositeľmi subjektívnych postojov k téme. Väčšinou sa jedná o prídavné mená (dobrý, zlý), menej sa vyskytujú podstatné mená (bomba) a slovesá (zničiť). Významnú úlohu hrajú aj príslovky, ktoré zvyšujú alebo znižujú intenzitu polaritu slova. Preto je nutné nájsť v texte slová, ktoré majú, respektíve nemajú význam pri klasifikácii názorov.

### A. Subjektivita slova

Určenie subjektivity slov znamená roztriedenie slov na tie, ktoré sú použiteľné pre ďalšiu klasifikáciu a ostatné, ktoré pre ďalšiu klasifikáciu potrebné nie sú. Aj keď sú sledované primárne prídavné mená, nie všetky prídavné mená sú nositeľmi potrebnej informácie pre určenie polaritu príspevku. To isté platí aj pre ostatné slovné druhy, ako sú napríklad podstatné mená, slovesá alebo príslovky.

Tabuľka 1  
 Roztriedenie slov na pozitívne, negatívne a neutrálne.

Pozitívne slová	Negatívne slová	Neutrálne slová
pecka, čarovný	blbý, nekvalitný	obyčajný, hlučne
očariť, dokonalý	des, zničiť	ateizmus, kresliť

Tabuľka 2  
 Roztriedenie slov podľa stupňov od silno pozitívnych až po silno negatívne.

Stupeň	Číselná hodnota	slová
silno pozitívne	3	veleúžasný, fenomenálny, legendárny, najkrajší
mierne pozitívne	2	úctyhodný, chváľitebný, taktný, pútavý
slabo pozitívne	1	odborný, dobrý, podarený, neškodný
slabo negatívne	-1	divý, zlý, biedny, čudný
mierne negatívne	-2	bizarný, desivý, chladný, nefudský
silno negatívne	-3	ohavný, tyranský, ohyzdny, zhnuseny

### B. Orientácia (polarita) slova

Poznáme tri základné stupne orientácie slov. Pozitívnu, negatívnu a neutrálnu polaritu. Podľa tohto delenia môžeme aj vyššie gramatické jednotky (vety a príspevky) rozdeliť na pozitívne, negatívne alebo neutrálne. V tabuľke 1 je možné vidieť rozdelenie práve na pozitívne, negatívne a neutrálne slová. Ak je to potrebné, je možné zvoliť aj jemnejšie delenie.

Existuje ale aj veľa slov, kde polaritu určuje kontext, v akom sa vo vete nachádzajú. Jedno slovo môže byť v jednej vete pozitívne, zatiaľ čo v druhej negatívne. Napríklad slovo *dlhý* je v prvej vete pozitívne a v druhej negatívne. [1]

*Batéria vo fotoaparáte vydrží dlhý čas.* – pozitívne

*Film bol veľmi dlhý.* - negatívne

Do úvahy treba brať aj ustálené slovné spojenia, ktoré majú ako celok určitú polaritu (Aj duša by vzal človeku, keby mu nedal pálenky.). Iným prípadom sú slovné spojenia s pozitívnymi alebo negatívnymi slovami, ktoré ale nenesú nijakú polaritu. Tu sa jedná najmä o anglické výrazy napríklad *pretty large*, kde *pretty* je pozitívne slovo, ale v danom spojení nemá žiadnu polaritu, alebo je jeho polarita daná kontextom, v akom sa v texte nachádza. [1]

*Negácia:* Problémom pri určovaní polarity slova môže byť aj otáčanie polarity slova pomocou záporu. Často sa vyskytujú vety, kde slovu s určitou polaritou predchádza zápor.

Príklad: *Nie je to z jeho strany pekné.*

Tento typ negácie je označovaný ako *switch negation*, kde sa význam slova mení na slovo s rovnakou silou ale opačnej polarite. Aj keď sa zdá, že tento typ negácie funguje, nie každé otočenie polarite funguje dobre. Problém nastáva pri otáčaní polarite slov s veľkou silou polarite. Keď budeme negovať slovo *najlepší* s hodnotou +3, logicky dostaneme ekvivalent slova *najhorší* s hodnotou -3. V skutočnosti *nie najlepší* s hodnotou -3 neoznačuje horšie hodnotenie ako *nie dobrý* s hodnotou -1. Preto bola zavedená aj *negácia posunom*(shift negation). Pri negácii posunom sa hodnota slova neurčuje otočením jeho hodnoty, ale posunom smerom k opačnej polarite o presnú sumu (napríklad o hodnotu 2). Takto sa pri slovách, ktorých sila orientácie je malá, polarita otáča, zatiaľ čo pri slovách, ktorých sila orientácie je vysoká, sa nová sila získava posunutím o určitú sumu. [3]

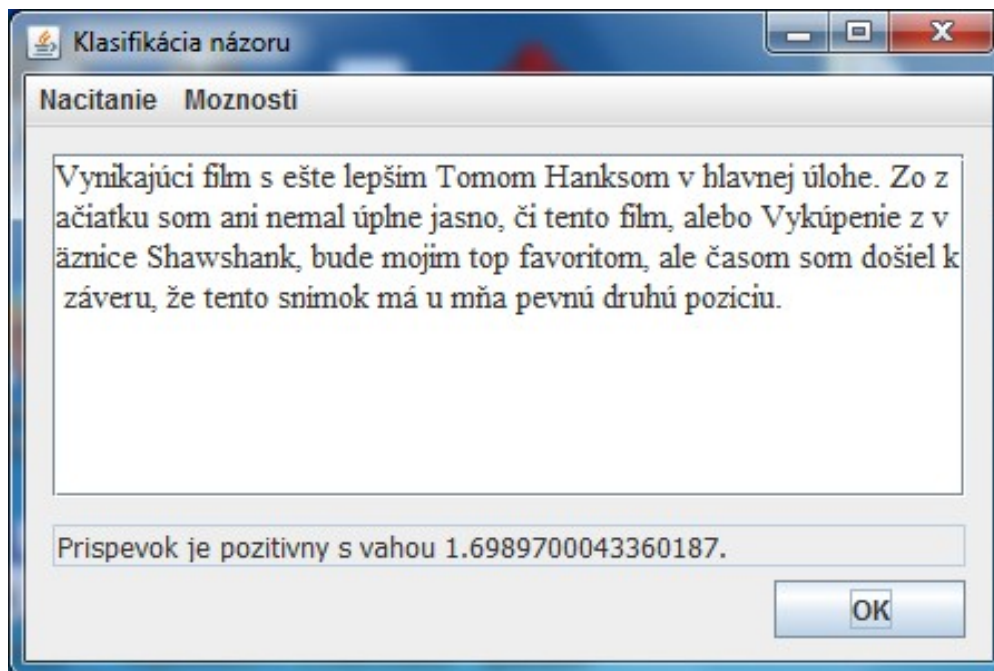
### C. Sila orientácie slova

Sila orientácie je sila, ktorou slovo, veta alebo diskusný príspevok potvrdzujú alebo negujú názor na tému, ktorá diskusiu vyvolala. Každé slovo má priradenú hodnotu na základe intenzity orientácie. Na základe takejto stupnice môžeme rozdeliť slová do skupín s príslušnými hodnotami.

Sila orientácie môže byť vyjadrená slovne, ale aj číselne. Pre ďalšie spracovanie je jednoduchšie používať číselné vyjadrenie. Keď budeme slová deliť od silno pozitívnych slov až po silno negatívne potom môžeme rozdeliť slová do skupín tak, ako je to v tabuľke 2.

*Intenzifikácia:* Pre zvýšenie presnosti klasifikácie je nutné sa na diskusné príspevky pozeráť zo širšieho pohľadu ako len na jednotlivé slová. Mali by sa brať do úvahy minimálne dvojice slov. Príspevky často obsahujú kombinácie slov, ktoré svojím významom posúvajú silu orientácie do vyššej, alebo nižšej roviny. Z tohto dôvodu by sa dali intenzifikátory rozdeliť na dve skupiny. Intenzifikátory, ktoré silu polarite zvyšujú, a tie, ktoré ju znižujú.

V niektorých prácach sa intenzifikátory používajú tak, že zvyšujú základnú silu o +1 alebo ju znižujú o -1. Toto hodnotenie však nezohľadňuje široké spektrum, v ktorom intenzifikátory



Obr. 1 Hlavné okno aplikácie na klasifikáciu názorov s príkladom komentára a jeho vyhodnotením.

pracujú. Napríklad slovo *extrémne* má väčšiu silu intenzifikácie ako slovo *celkom*. Ďalším problémom je, že zosilnenie už aj tak silného výrazu je väčšie, ako zosilnenie slabšieho výrazu pomocou rovnakého intenzifikátora (porovnanie *extrémne fantastický* s *extrémne dobrý*). Preto sa ako najvhodnejšie riešenie javí intenzifikácia pomocou percent. [3]

#### D. Tvorba slovníkov

Pre potreby klasifikácie názorov je treba vytvoriť slovník, v ktorom na nachádzajú všetky slová, ktoré sú nositeľmi názoru v danej oblasti alebo doméne. Veľmi málo ľudí vyjadrujúcich sa na webe používa spisovný jazyk. Je potrebné, aby sa tomuto javu prispôbil aj slovník. Slovník by teda okrem spisovných výrazov mal obsahovať aj slangové alebo nespisovné slová často používané práve mladou generáciou.

### III. IMPLEMENTÁCIA A TESTOVANIE ALGORITMU NA KLASIFIKÁCIU NÁZOROV

Bola vytvorená aplikácia, ktorá slúži na klasifikáciu názorov dolovaných v konverzačnom obsahu. Príspevky je možné zadávať manuálne, automaticky zo súboru alebo z webu. Program využíva slovník uložený v databáze. Výstupom aplikácie je identifikácia intenzity polarity komentárov a následne aj celej diskusie. Aplikácia je implementovaná v jazyku JAVA, teda je možné spustiť ju na akomkoľvek stroji, na ktorom sa nachádza JVM (JAVA Virtual Machine). V hlavnom okne môžeme prepínať medzi rôznymi funkciami programu. V režimoch načítania diskusie je možné načítať diskusie z rôznych zdrojov. V časti možnosti môže používateľ nastavovať pripojenie k databáze, upravovať slovník, webové stránky a k nim prisluchajúce tagy. Na obrázku 1 je zobrazené hlavné okno aplikácie.

Súčasťou programu je aj slovník, v ktorom sú kľúčové slová potrebné pre klasifikáciu. Prvý slovník bol vytvorený prekladom z anglického slovníka, ktorý obsahoval pozitívne a negatívne slová. Následne boli nájdené všetky synonymá preložených slov. Ďalšie slovníky boli vytvorené analýzou chybných výsledkov, kde do slovníka boli pridané intenzifikátory a negátory.

Ak slová v slovníku končia samohláskou, okrem *o*, sú zapísané v tvare nominatívu množného čísla. Napríklad slovo *dobrý* je v slovníku v tvare *dobri*. V prípade, že slovo končí spoluhláskou, alebo samohláskou *o*, je zapísané v tvare bez úprav. Každému slovu v slovníku je priradená polarita. Pozitívna polarita je označená ako *p*, negatívna je označená ako *n*. Intenzifikátory majú označenie *i* a záporné *o*. Slovám sme na základe subjektívneho hodnotenia priradili aj silu orientácie na stupnici od -3 po 3. Slovám, ktoré otáčajú polaritu je pridelená nula, lebo algoritmus v prípade, že identifikuje zápor, automaticky otočí polaritu.

poz: 1161	neutr: 94	neg: 84
pus: 1029	pun: 132	
nus: 44	nun: 40	
neus: 68	neun: 26	
poc_poz: 1083	poc_neg: 95	poc_neu: 161
pres poz: 95	návr poz: 88.6	
pres neg: 46.3	návr neg: 52.4	
pres neu: 42.2	návr neu: 72.3	

Obr. 2 Výsledok testovania pre film Zelená míľa za použitia slovníka so všetkými typmi slov

poz: 1169	neutr: 147	neg: 94
pus: 979	pun: 190	
nus: 30	nun: 64	
nues: 83	neun: 64	
poc_poz: 1078	poc_neg: 85	poc_neu: 247
pres poz: 90.8	návr poz: 86.3	
pres neg: 39.7	návr neg: 28.7	
pres neu: 36.4	návr neu: 56.5	

Obr. 3 Výsledok testovania pre film Forrest Gump za použitia slovníka so všetkými typmi slov

#### A. Testovanie riešenia

Ako testovacie dáta boli vybrané časti diskusií k filmom Zelená míľa a Forrest Gump z webovej stránky [www.csf.d.cz](http://www.csf.d.cz). Tieto príspevky boli preložené do slovenského jazyka a kvôli objektívite v nich boli ponechané všetky gramatické a štylistické chyby. Preto by mali prezentovať reálne situácie v konverzácii alebo hodnotenie filmov respektívne produktov. Príspevkov hodnotiacich film Zelená míľa bolo 1339 a Forrest Gump 1410. Celkovo má teda testovaná vzorka 2749 príspevkov.

*Testovanie diskusie k filmu Zelená míľa:* Diskusia k filmu Zelená míľa obsahuje 1339 komentárov. Diskusia obsahuje 1161 pozitívnych, 84 negatívnych a 94 neutrálnych komentárov. Na obrázku 2 môžeme vidieť vyhodnotenie tohto testu. Z obrázka vyčítame, že program našiel 1083/1161 komentárov s pozitívnym hodnotením, z čoho sa s expertom zhodli v 1029 prípadoch. Negatívnych príspevkov bolo identifikovaných 95/84, z čoho bolo 44 klasifikovaných správne. Počet neutrálnych príspevkov objavených v diskusii bolo 161/94 a dobre určených bolo 68. Program nesprávne určil 132 pozitívnych, 40 negatívnych a 26 neutrálnych príspevkov. Presnosť pri určovaní pozitívnych príspevkov bola 95% a návratnosť 88.6%. Negatívne komentáre mali presnosť 46.3% a návratnosť 52.4%. Nakoniec presnosť pre neutrálne príspevky bola 42.2% a návratnosť 72.3%.

*Testovanie diskusie k filmu Forrest Gump:* Diskusia k filmu Forrest Gump obsahuje 1410 komentárov. Diskusia obsahuje 1169 pozitívnych, 94 negatívnych a 147 neutrálnych komentárov. Na obrázku 3 môžeme vidieť vyhodnotenie tohto testu, z čoho je zjavné, že program identifikoval 1078/1169 pozitívnych príspevkov. S expertom sa zhodli v 979 prípadoch. Negatívnych komentárov bolo nájdených 85/94, z tohto počtu bolo korektne zatriedených 30 príspevkov. Neutrálne označených komentárov bolo 247/147. Z nich bolo 83 zaradených dobre. Program nesprávne určil 190 pozitívnych, 64 negatívnych a 64 neutrálnych príspevkov. Presnosť pri určovaní pozitívnych príspevkov bola 90.8% a návratnosť 86.3%. Negatívne komentáre mali presnosť 39.7% a návratnosť 28.7%. Nakoniec presnosť pre neutrálne príspevky bola 36.4% a návratnosť 56.5%.

Na prezentovaných výsledkoch má zásluhu veľké množstvo pozitívnych príspevkov. Testovacie dáta boli vybrané tak, aby obsahovali veľa príspevkov k jednej téme. Preto boli vybrané príspevky k veľmi vysoko a kladne hodnoteným filmom. To sa podpísalo na množstve pozitívnych komentárov a malom počte negatívnych. Ako neutrálne príspevky boli väčšinou hodnotené príspevky, ktoré hodnotili objektívnu stránku filmu, ako napríklad výroky z filmu, komentáre z natáčania alebo nominácie na Oscarov.

Často sa k nepresnostiam pripájali aj komentáre, v ktorých autor najskôr opisoval filmy ako



pozitívne, ale čím viackrát ich videl, tým menej sa mu páčili. Komentár bol teda vyhodnotený expertom ako negatívny, ale program ho na základe kladného počiatočného opisu vyhodnotil pozitívne. Takisto, keď sa hodnotil film pozitívne, jednalo sa väčšinou o priame hodnotenie, zatiaľ čo negatívne hodnotenie vychádzalo skôr z opisov častí, ktoré sa autorovi komentáru nepáčili. Ďalším typom komentárov, ktoré robili problémy boli také, ktoré hodnotili film len bodovým alebo percentuálnym hodnotením bez, alebo s minimálnym dodatočným opisom. Posledným problémom je identifikácia irónie. Príspevky často obsahovali kombinácie neutrálnych slov, ktoré mali spolu pozitívny alebo negatívny podtext.

#### IV. ZÁVER

Programy analyzujúce webové diskusie budú čím ďalej tým populárnejšie, lebo pri rastúcom počte diskutujúcich a príspevkov bude čoraz náročnejšie prečítať celé diskusie. V tejto práci bol vytvorený nový slovníkový prístup k identifikácii názorovej polaroty webovej diskusie. Aplikácia klasifikuje slová, kombinácie slov, až po úroveň príspevku a celej diskusie. V algoritme je implementovaná možnosť spracovania intenzifikácie negácie a aj negácie intenzifikácie poprípade spracovanie viacnásobnej intenzifikácie alebo negácie. Uvedené riešenie by malo poskytovať čo najvernejšie spracovanie bežného jazyka.

Hodnoty presnosti a návratnosti pre pozitívne príspevky možno považovať ako dobrý výsledok. Ostatné hodnoty boli ovplyvnené najmä nízkym počtom príspevkov. V budúcnosti by bolo dobré otestovať algoritmus na diskusií s rovnomernejším rozložením príspevkov. Ďalším výrazným problémom bola identifikácia irónie a dvojsmyselnosti.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in *Proceedings of the 2008 International Conference on Web Search and Data Mining*, ser. WSDM '08. New York, NY, USA: ACM, 2008, pp. 231–240.
- [2] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.
- [3] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Comput. Linguist.*, vol. 37, no. 2, pp. 267–307, jun 2011.

# Kodér a dekodér stavového binárneho aritmetického kódovania

Iveta GLADIŠOVÁ

Katedra elektroniky a multimediálnych telekomunikácií  
Fakulta elektrotechniky a informatiky, Technická univerzita Košice, SR

iveta.gladisova@tuke.sk

**Abstrakt** — Článok sa zaoberá binárnym aritmetickým kódovaním a to predovšetkým stavovým a stavovým adaptívnym. Popisuje vo všeobecnosti rôzne modifikácie binárneho aritmetického kódovania. Experimentálna časť vyhodnocuje dosiahnuté výsledky simulácie stavového binárneho aritmetického kodéra a stavového adaptívneho binárneho aritmetického kodéra z hľadiska použitia rôznych šablón na skupine testovaných binárnych obrazoch a výsledné dĺžky kódov, ako aj kompresné pomery pri takomto kódovaní binárnych rovín obrazu.

**Kľúčové slová** — aritmetické kódovanie, JBIC šablóna, stavové aritmetické kódovanie, bitové roviny .

## I. ÚVOD

Aritmetické kódovanie (AK) patri medzi nerovnomerné tzv. entropické kódy, ktoré efektívne redukovujú štatistickú redundanciu, reprezentovanú nelineárnosťou rozdelenia pravdepodobnosti symbolov, obsiahnutú vo vstupných dátach [1], [8]. AK obchádza myšlienku nahradenia symbolu špecifickým kódom s celistvým počtom bitov a namiesto toho nahrádza tok vstupných symbolov jednoduchým desatinným číslom z polootvoreného intervalu pravdepodobností  $< 0, 1$ ). Nevýhodou aritmetického kódu je veľký počet operácií potrebných pre proces kódovania, avšak dosahuje vysokú účinnosť kódovania. V aritmetickom kódovacom systéme musí kodér spolupracovať s modelom, ktorý poskytuje informáciu o pravdepodobnosti. Kodér potom využíva tieto pravdepodobnosti na zakódovanie aktuálneho symbolu. Na zabezpečenie dekódovateľnosti môže model použiť len informácie známe kodéru aj dekodéru. Model sa môže zmeniť ihneď po zakódovaní vstupného súboru [3], [4], [6]:

## II. BINÁRNE ARITMETICKÉ KÓDOVANIE

### *Neadaptívny binárny aritmetický kodér*

Neadaptívny binárny AK (NBAK) patrí svojim algoritmom do skupiny jednoduchšie technicky realizovateľných binárnych AK. Pre štatistickú analýzu obrazu využíva stacionárny estimátor pravdepodobnosti 0. rádu. Jeho algoritmus je nasledovný [8], [9]:

1. Estimácia pravdepodobnosti výskytu, t.j. pre všetky prvky obrazu sa vykoná cyklus:
  - a) určí sa aktuálny symbol,
  - b) inkrementuje sa príslušné počítadlo symbolu.
2. Porovnaním počítadiel sa určí LPS (less probable symbol -menej pravdepodobný symbol).
3. Určí sa estimovaná pravdepodobnosť LPS.
4. Vykoná sa aritmetické kódovanie.

NBAK dosahuje v priemere nižšiu účinnosť kompresie, čo je spôsobené rádom estimátora pravdepodobnosti. Spomedzi spomínaných variácií kodérov má najnižšie výpočtové nároky a je teda najrýchlejší .

### *Adaptívny binárny aritmetický kodér*

Adaptívny binárny AK (ABAK) - štatistická analýza obrazu sa vykonáva pomocou adaptívneho estimátora pravdepodobnosti 0. rádu. Jeho algoritmus je nasledovný [7].:

Pre všetky prvky obrazu sa vykoná cyklus:

- a) určí sa aktuálny symbol,
- b) určí sa LPS a jeho estimovaná pravdepodobnosť pre minulý krok cyklu,
- c) inkrementuje sa príslušné počítadlo symbolu,
- d) vykoná sa aritmetické kódovanie.

V každom z týchto cyklov sa menia podmienky, za ktorých dochádza ku kódovaniu aktuálneho symbolu, čo je ekvivalentné aritmetickému kódovaniu s cyklickou estimáciou pravdepodobnosti na začiatku. Estimácia pravdepodobnosti je vykonávaná na báze znalosti štatistiky dosiaľ kódovaných symbolov. Aby bolo možné uskutočniť v dekodéri totožný cyklus, musí byť splnená nutná podmienka, vyplývajúca z kroku b). Na zakódovanie postupnosti je potrebná inicializácia LPS a jeho pravdepodobnosti, ktorá prebieha iba raz, a to pred začiatkom samotného kódovania. V priemere nižšiu účinnosť kompresie zapríčiňuje najmä rád estimátora a adaptívna estimácia. Výpočtové nároky sú vzhľadom na opakovanú estimáciu pravdepodobnosti vyššie, rýchlosť výpočtu tak klesá s počtom obrazových prvkov kódovaného obrazu.

#### **Stavový binárny aritmetický kodér**

Realizácia stavových binárnych aritmetických kodérov (SBAK) patrí medzi zložitejšie – mení sa postup estimácie pravdepodobnosti a aritmetické kódovanie je nahradené stavovým aritmetickým kódovaním [3], [10]. Štatistická analýza obrazu je vykonávaná stacionárnym estimátorom pravdepodobnosti  $n$ -tého rádu, rád estimátora určuje užívateľom zvolená šablóna.

Jeho algoritmus je nasledovný :

1. Estimácia pravdepodobnosti – pre všetky prvky obrazu sa vykoná cyklus:
  - a) určí sa aktuálny stav daný šablónou a aktuálny symbol v šablóne,
  - b) inkrementuje sa príslušné počítadlo stavu pre aktuálny symbol.
2. Porovnaním počítadiel sa pre jednotlivé stavy určí LPS v rámci stavu a jemu odpovedajúcu estimovanú pravdepodobnosť.
3. Stavové aritmetické kódovanie – pre všetky prvky obrazu sa vykoná cyklus :
  - a) pre aktuálny prvok sa určí aktuálny stav,
  - b) pre aktuálny stav sa z modelu zistí LPS a jeho estimovaná pravdepodobnosť,
  - c) vykoná sa aritmetické kódovanie pre aktuálny prvok.

SBAK dosahuje vyššiu účinnosť kompresie vďaka vhodne zvolenému estimátoru. Je nutné poznamenať, že dosiahnutá účinnosť kompresie sa bude líšiť v závislosti od zvolenej šablóny. V priemere dosahovaná účinnosť kompresie je však vyššia, než v prípade nestavových kodérov. Kladené sú však i väčšie výpočtové nároky - nielen na procesor, ale i na pamäť, ktorá rastie exponenciálne so zvoleným rádom šablóny. Rýchlosť výpočtu výrazne klesá so vzrastom počtu prvkov šablóny i počtom obrazových prvkov kódovaného obrazu.

Spomedzi spomínaných AK je pre tento typ kodéra nutné vysielat' najviac extra-dát, konkrétne, okrem rozmeru obrazu je to i použitý model. Výstupný kód sa tak môže zväčšiť rádovo o niekoľko stoviek bajtov [4], [2].

#### **Stavový adaptívny binárny aritmetický kodér**

Realizácia stavového adaptívneho binárneho AK (SABAK) je z pohľadu realizácie kombináciou stavového a adaptívneho AK. Štatistická analýza obrazu je tu vykonávaná adaptívnym estimátorom pravdepodobnosti  $n$ -tého rádu, rád estimátora je určený užívateľom zvolenou šablónou. Jeho algoritmus je nasledovný: pre všetky prvky obrazu sa vykoná cyklus:

- a) určí sa aktuálny stav daný šablónou a aktuálny symbol v šablóne,
- b) vzorcom sa určí LPS a jemu prislúchajúca estimovaná pravdepodobnosť aktuálneho stavu pre minulý krok cyklu,
- c) inkrementuje sa príslušné počítadlo stavu pre aktuálny symbol,
- d) vykoná sa aritmetické kódovanie

Estimácia pravdepodobnosti sa vykonáva pre daný stav v každom cykle kódovania (cyklom sa rozumie kódovanie jedného obrazového prvku), jej výsledok sa však použije až pri ďalšom výskyte tohto stavu. Takéto prevedenie estimátora vyžaduje jeho inicializáciu, zabezpečuje však dekódovateľnosť zakódovaného obrazu.

SABAK dosahuje vyššiu účinnosť kompresie vďaka vhodne zvolenému estimátoru. Účinnosť kompresie závisí na voľbe šablóny a rýchlosti adaptácie štatistických parametrov (pravdepodobnosti). SABAK má zvýšené výpočtové nároky – v každom cykle prebieha množstvo operácií naraz. Väčšia pamäť je potrebná iba v čase výpočtov. Rýchlosť výpočtu výrazne klesá so vzrastom počtu prvkov šablóny a s počtom obrazových prvkov kódovaného obrazu [10], [5], [6].

SABAK v sebe implementuje stavové a adaptívne kódovanie, z čoho vyplýva i systém detekcie aktuálneho stavu a estimácie aktuálnej pravdepodobnosti LPS. Detekcia aktuálneho stavu a estimácia pravdepodobnosti funguje nasledovne [10]:

$$Ak n(0) > n(1) \Rightarrow LPS = "1" ; \text{ inak : } LPS = "0" \quad (1)$$

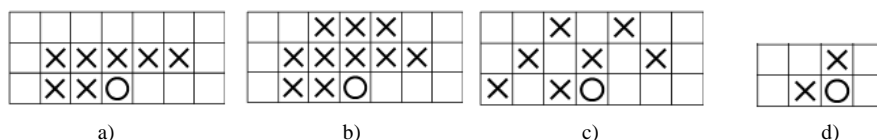
$$p(LPS) = \frac{n(LPS) + 1}{n(LPS) + n(MPS) + 2} , \text{ kde} \quad (2)$$

- $n(0)$  - početnosť aktuálneho stavu so symbolom "0" v šablóne pre daný cyklus
- $n(1)$  - početnosť aktuálneho stavu so symbolom "1" v šablóne pre daný cyklus
- $p(LPS)$  - estimovaná pravdepodobnosť LPS
- $n(LPS)$  - početnosť aktuálneho stavu so symbolom. LPS v šablóne pre daný cyklus
- $n(MPS)$  - početnosť aktuálneho stavu so symbolom MPS (most probably symbol) v šablóne pre daný cyklus.

### Model SABAK

Pojmom model súhrnne označujeme tú časť informácií v rámci extra-dát, ktorá určuje počet prvkov šablóny, ich usporiadanie v dvojrozmernom priestore a pre ňou určené stavy hodnôt Q a LPS [3], [10]. Model tak plní úlohu knižnice dát, zaznamenávajúcej dáta potrebné pri estimácii a aproximácii pravdepodobnosti, potrebné jak pre kodér tak i pre dekodér.

Pre binárne obrazy je model charakterizovaný šablónou, určujúcou rád a usporiadanie bodov obrazu voči aktuálnemu bodu estimátora pravdepodobnosti. Jej tvar je navrhovaný za účelom dosiahnutia čo najlepšej účinnosti kompresie obrazu, musí však spĺňať podmienku kauzality, ktorú je možné interpretovať nasledovne: estimácia pravdepodobnosti aktuálneho symbolu musí prebiehať iba na základe už známych obrazových bodov, teda všetky body šablóny sa musia nachádzať pred alebo nad aktuálnym bodom (krúžok na Obr.1). Na Obr.1 sú niektoré známe typy používaných šablón a), b)[3] a tiež typy šablón c), d), ktoré boli použité pri simuláciách, nižšie v článku.



Obr.1 Tvary šablón a) ABIC, b) JBIG, c) náhodná, d) jednoduchá.

Stavový estimátor pravdepodobnosti funguje na princípe preloženia danej šablóny nad binárny obraz v mieste aktuálneho bodu a pre tento bod určí jeho stav. Stav je jednoznačne určený hodnotami vo všetkých bodoch šablóny pomocou všeobecného vzorca [3], [10]:

$$S(y) = \sum_{i=0}^{N-1} c_i \cdot 2^i = c_0 + 2 \cdot c_1 + 4 \cdot c_2 + 8 \cdot c_3 + \dots, \text{ kde} \quad (3)$$

$S(y)$  - hodnota stavu pre aktuálnu pozíciu šablóny

$c_i$  - hodnota v  $i$ -tom bode šablóny

$N$  - počet prvkov šablóny.

Určenie modelu binárneho obrazu spočíva v zistení počtu výskytov jednotlivých stavov a im prislúchajúcich aktuálnych obrazových prvkov v obraze. Pre každý stav teda určíme podmienenú pravdepodobnosť stavu v závislosti od aktuálneho bodu obrazu, pričom výsledné pravdepodobnosti budú tvoriť model. Podmienené pravdepodobnosti určíme nasledovným spôsobom [3], [10]:

$$p(c_A = v / S_k) = p(S_{k,v}) = \frac{n_{k,v}}{n_k}, \text{ kde} \quad (4)$$

$c_A$  - všeobecná hodnota aktuálneho obrazového prvku šablóny

$v$  - hodnota obrazového prvku - "0" alebo "1"

$S_k$  -  $k$ -tý stav (počet stavov šablóny je  $2^N$ )

$S_{k,v}$  -  $k$ -tý stav s aktuálnym obrazovým prvkom "v"

$p(S_{k,v})$  - podmienená pravdepodobnosť  $k$ -tého stavu s aktuálnym obrazovým prvkom "v"

$n_k$  - početnosť  $k$ -tého stavu (ak je rovný nule, definujeme  $p(S_{k,v}) = 0$ )

$n_{k,v}$  - početnosť  $k$ -tého stavu s aktuálnym obrazovým prvkom "v".

Model SABAK je vytváraný v procese kódovania postupnou adaptáciou, jeho veľkosť je určená rozmerom resp. počtom prvkov šablóny. Tvar šablóny nie je vopred určený, volí sa s ohľadom na optimálne výsledky pri kódovaní binárneho obrazu. Inicializácia modelu SABAK sa uskutoční nastavením  $S_{k,v} = 0$  pre jednotlivé stavy a aktualizácia modelu prebieha počas procesu kódovania v zmysle rovníc (1) a (2). V každom kroku sa pritom model prispôsobí výsledkom detekcie nového obrazového prvku a určenia jemu prislúchajúceho stavu. Prechodom cez celý obraz získame úplný model, ktorý však nie je potrebné v úplnom znení vyslať dekodéru [10], [5].

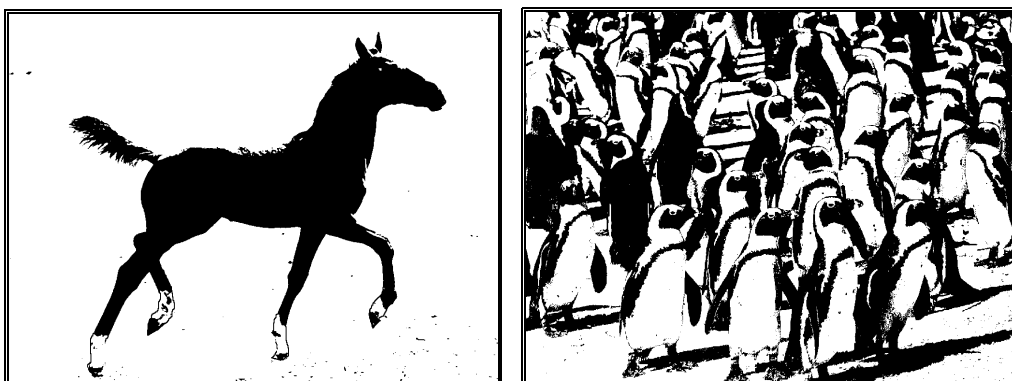
### III. DOSIAHNUTÉ VÝSLEDKY SIMULÁCIE SBAK A SABAK

Metódy stavového binárneho aritmetického kódovania a stavového adaptívneho binárneho aritmetického kódovania boli simulované navrhnutým programom v programovom jazyku Builder C++. Skúmala sa účinnosť SBAK a SABAK na jednotlivé obrázky a sledovalo sa, ako sa bude meniť výsledná kompresia v závislosti na použitej šablóne. Kodérmi SBAK a SABAK sa tiež zakódoval obraz Lena rozdelený do bitových rovín, ako je to uvedené v [4].

Binárne testovacie obrázky boli vybrané podľa: rozmeru a zložitosti (detailnosti). Použili sa obrázky malej veľkosti 256x256 obrazových prvkov a veľkosti 1024x768. Podľa zložitosti obrazov sa vybrali jednoduché obrázky (málo detailov) a zložité obrázky (veľa detailov) (pri výbere obrazov si treba uvedomiť, že detailnosť obrazu bola určovaná subjektívne). Jednotlivé obrázky, ktoré boli použité pri simulácii BAK sú na Obr.2 a Obr.3.



Obr.2 Binárne obrázky veľkosti 256x256 obrazových prvkov s názvami „ruza“ (vľavo), „papagaj“ (stred) a „slnecnice“ (vpravo).



Obr.3 Binárne obrázky veľkosti 1024x768 obrazových prvkov s názvami „kon“ (vľavo) a „tucniaky“ (vpravo).

Skúmala sa najprv účinnosť SABAK pre rôzne šablóny (JBIG, ABIC, jednoduchá a náhodná – vid' obr.1). Na vyjadrenie účinnosti kompresie sa používa kompresný pomer, ktorý možno vyjadriť v tvare

$$C_r = \frac{N_k}{N}, \quad (5)$$

kde  $N$  je počet jednotiek (bitov, bytov) nezakódovanej správy a  $N_k$  je počet jednotiek zakódovanej správy. Pre efektívne kódovanie musí platiť, že  $C_r < 1$ . S nárastom hodnoty kompresného pomeru klesá účinnosť kódovania.

Hodnoty kompresného pomeru pri SABAK a pre jednotlivé šablóny sú v Tab 1.

Tab.1 Kompresné pomery SABAK pre skúšobné obrázky

obraz šablóna	ruza	papagaj	slnecnice	kon	tucniaky
JBIG	<b>0,033</b>	0,137	0,508	<b>0,02</b>	<b>0,136</b>
ABIC	0,035	<b>0,136</b>	<b>0,5</b>	0,021	0,138
jednoduchá	0,052	0,158	0,545	0,032	0,166
náhodná	0,036	0,139	0,51	0,021	0,141

Z výsledkov uvedených v Tab.1 vyplynulo:

- ako najuniverzálnejšia šablóna pre SABAK je JBIG (potvrdenie teoretic. predpokladu),
- pre obrázky s malými detailmi bol dosiahnutý najlepší kompresný pomer,
- najhorší kompresný bol dosiahnutý pri obrazoch s veľa detailmi,
- JBIG a ABIC šablóny dosahovali porovnateľné výsledky, kde rozdiel pri kompresnom pomere bol až na 3 významovom mieste,

- s nárastom rozmeru obrazov pri obrazoch s veľa detailmi sa znížil kompresný pomer (tucniaky a slnečnice), z čoho vyplýva, že SABAK je menej vhodný pre obrazy malých rozmerov a s veľkým množstvom detailov,
- pri malom ráde šablóny (jednoduchá) boli dosiahnuté najhoršie kompresné pomery.

V Tab. 2 sú uvedené výsledky dosiahnutých kompresných pomerov pre SABAK, pričom sa použili uvedené skúšobné obrazy a šablóny.

Tab.2 Kompresné pomery SABAK pre skúšobné obrazy

obraz šablóna	ruza	papagaj	slnečnica	kon	tucniaky
JBIG	0,1042	0,1965	0,5505	0,0253	0,1389
ABIC	<b>0,0412</b>	<b>0,1396</b>	<b>0,5018</b>	<b>0,0213</b>	<b>0,1379</b>
jednoduchá	0,0500	0,1562	0,5435	0,0317	0,1660
náhodná	0,0413	0,1413	0,5105	<b>0,0213</b>	0,1409

Porovnaním výsledkov z tabuliek 1 a 2 vyplynulo:

- SABAK bol pri všetkých obrazoch účinnejší
- pri SABAK sa javí ako najuniverzálnejšia šablóna ABIC, zatiaľ čo pri SABAK je to JBIG šablóna
- stavová adaptácia SABAK sa najviac prejavila pri obrazoch s málo detailmi

Pri simulácii bitových rovín boli použité už len šablóny typu JBIG a ABIC. Jednotlivé bitové roviny sú prevzaté z [4] a v tabuľkách 3 a 4 znamená BR 0 bitovú rovinu 0.váhy (vytvorená s LSB), BR 1 je bitová rovina 1.váhy atď. V Tab.3 a Tab.4 sú uvádzané veľkosti zakódovaných postupností v bitoch. Nezakódovaný obraz je vo všetkých prípadoch veľkosti 65536 bitov (256 x 256). Ako hlavný ukazovateľ slúži celkový kompresný pomer pri „Suma BR 0 až BR 7“.

Tab. 3 Kompresné pomery SABAK pre bitové roviny

Bitová rovina	Zakódované pomocou SABAK a šablónou typu		Kompresný pomer	
	JBIG [bit]	ABIC [bit]	JBIG	ABIC
BR 0	70587	66177	1,077	1,009
BR 1	70490	66176	1,076	1,009
BR 2	69689	65846	1,064	1,005
BR 3	61276	58452	1,069	0,892
BR 4	49913	46831	0,762	0,715
BR 5	34049	31415	0,519	0,479
BR 6	25044	22047	0,382	0,336
BR 7	15094	11767	0,23	0,18
		Suma BR 0 až BR 7	6,179	5,625

Tab. 4 Kompresné pomery SABAK pre bitové roviny

Bitová rovina	Zakódované pomocou SABAK a šablónou typu		Kompresný pomer	
	JBIG [bit]	ABIC [bit]	JBIG	ABIC
BR 0	68007	66013	1,038	1,007
BR 1	68001	66025	1,038	1,007
BR 2	67339	65715	1,025	1,003
BR 3	58935	58447	0,899	0,892
BR 4	46759	46787	0,713	0,714
BR 5	31113	31341	0,475	0,478
BR 6	21889	21885	0,334	0,334
BR 7	11491	11585	0,175	0,177
		Suma BR 0 až BR 7	5,697	5,612

Z porovnania dosiahnutých výsledkov pre SBAK a SABAK bitových rovín v Tab. 3 a 4 vyplynulo:

- SABAK dosiahol lepší výsledok kompresného pomeru o 0,482 oproti SBAK pri použití JBIG šablóny, čo znova potvrdilo, že SABAK je účinnejší oproti SBAK,
- SABAK dosiahol pri použití ABIC šablóny lepší výsledok ako SBAK, ale rozdiel je minimálny,
- z porovnania účinnosti kódovania SBAK a SABAK pri jednotlivých bitových rovinách so šablónou JBIG vidno, že napríklad pri BR 0 je rozdiel minimálny (lebo táto bitová rovina má najnižší stupeň korelácie a jedná sa vlastne o šum) a pri BR 7 je SABAK znovu lepší (lebo táto bitová rovina má najvyšší stupeň korelácie). Pri šablóne ABIC sú rozdiely minimálne,
- vo všeobecnosti SABAK u viacúrovňových obrazov s rozkladom na bitové roviny dosahuje vysokú kompresiu, pričom nedochádza k žiadnej strate informácie tohto obrazu,
- pri SABAK bitových rovín sa využívajú korelačné väzby medzi ich obrazovými prvkami, keď so stúpajúcou váhou bitových rovín stúpa aj korelácia a tým je možné zvýšiť aj účinnosť kódovania. Napr. pre bitovú rovinu BR 7 s najvyššou koreláciou je najnižší kompresný pomer 0,175 oproti BR 0, kde je tento pomer 1,038.

Výsledky kódovania, ktoré porovnávali SBAK a SABAK, boli uvedené bez extra dát (to znamená, že k výslednej postupnosti bitov sa nezarátala pri SABAK šablóna a rozmer obrazu. Pri SBAK sa nezarátal ešte navyše použitý model).

#### IV. ZÁVER

Článok je zameraný na porovnanie kodérov stavového binárne aritmetického kódovania a stavového adaptívneho binárneho aritmetického kódovania. Popísal modifikácie binárneho aritmetického kódu a tiež jednotlivé princípy kodérov, ktoré sú založené na tomto binárnom aritmetickom kóde. Podrobnejšie sa venuje SABAK, kde sa uvádza, ako sa vytvára model pre SABAK. Dostupnými programovými prostriedkami bol zrealizovaný kodér a dekodér stavového binárneho aritmetického kódovania a otestovala sa účinnosť SBAK a SABAK na rôzne druhy obrazov s použitím rôznych šablón. Ďalej sme zisťovali účinnosť SABAK na obraze, ktorý bol rozdelený do jednotlivých bitových rovín. Tieto výsledky sa porovnali s výsledkami pre SBAK. Z výsledkov vyplynulo, že najuniverzálnejšia šablóna pre SABAK je JBIG šablóna, SABAK dosahuje najlepší kompresný pomer pre obrazy s malými detailmi, SABAK je menej vhodný pre obrazy malých rozmerov a s veľkým množstvom detailov. Výsledky, ktoré sa dosiahli, sa overili pomocou dekodéra, ktorý je tiež obsiahnutý v našich programových prostriedkoch.

#### LITERATÚRA

- [1] Mihalík, J., *Kódovanie obrazu vo videokomunikáciách*. Košice : Mercury-Smekál , 2001.
- [2] Marpe D. - Schwarz,H. - Wiegand, T., *Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard*. IEEE Trans. Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 620–636, July, 2003.
- [3] Gladišová, I. - Mihalík, J., *Stavové aritmetické kódovanie binárnych obrazov*. Košice, Acta Electrotechnica et Informatica, Vol.3, No.3, s.36-43, 2003.
- [4] Gladišová, I. - Mihalík, J. – Zavacký, J., *Bezstratová kompresia obrazu pomocou stavového binárneho aritmetického kódovania jeho bitových rovín*. Slaboproudý obzor, Vol.62, No.2, s.18-21, 2006.
- [5] Kim, Ch. H. - Park, I. Ch., *Parallel Decoding of Context-Based Adaptive Binary Arithmetic Codes Based on Most Probable Symbol Prediction*. IEICE TRANSACTIONS on Information and Systems, Vol.E90-D, No.2, pp.609-612, 2007.
- [6] Gladišová, I. - Mihalík, J., *Stavové aritmetické kódovanie binárneho tvaru videoobjektu*. Zborník ved. konf. s medzinárodnou účasťou "Nové smery v spracovaní signálov VI", Tatranské Zruby, s.219-222, 2002.
- [7] Marpe D., *Fast Adaptive Binary Arithmetic Coding (M Coder)*. <http://iphome.hhi.de/marpe/mcoder.htm> (active July 2014)
- [8] Witten, I. H.- Neal, R. M. - Cleary, J. G. :*Arithmetic Coding for Data Compression*. Communication of the ACM. Vol. 30, pp. 520-540, June 1987.
- [9] Langdom, G. G.- Rissanen, J. :*Compression of Black - White Images with Arithmetic Coding*. IEEE Transactions on Communications, Vol. COM-29, No. 6, pp. 858-867, 1981.
- [10] Marpe, D. – Blattermann, G. – Heising, G. – Wiegand, T. : *Video Compression Using Context-based Adaptive Arithmetic Coding*. Proc. IEEE ICIP, Vol. 3, pp.558-561, Oct.2001.

# Rozpoznávanie sarkastických viet v konverzačnom obsahu

<sup>1</sup>Richard LÖRINC, <sup>1</sup>Kristína MACHOVÁ

<sup>1</sup>Katedra Kybernetiky a Umelej inteligencie, Fakulta Elektrotechniky a Informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>richard.lorinc@student.tuke.sk, <sup>2</sup>kristina.machova@tuke.sk

**Abstract** — Sarkazmus je sofistikovaná forma reči používaná hlavne v on-line komunitách. Automatické rozpoznávanie sarkazmu je však nová úloha. Rozpoznávanie sarkazmu by mohlo zlepšiť presnosť analýzy názorov a postojov. V tomto príspevku je popísaný prehľad prístupov k dolovaniu konverzačného obsahu na diskusných fórach, na stránkach obsahujúcich príspevky na rôzne produkty a v rámci sociálnych sietí. Taktiež charakterizuje problémy analýzy sentimentu. V príspevku je navrhnutá metóda automatického rozpoznávania sarkazmu založená na identifikácii preddefinovaných vzorov. Jej cieľom je vyhodnotenie stupňa sarkazmu pomocou preddefinovaných vzorov na základe slovníkového prístupu. Vysvetľuje navrhnuté riešenie a popisuje jeho princíp. Na záver článok vyhodnocuje úspešnosť tejto metódy.

**KLúčové slová** — Sarkazmus, analýza názorov, dolovanie konverzačného obsahu, slovníkový prístup

## I. ÚVOD

Nepriama reč je sofistikovaná forma rozprávania v ktorej rozprávači oznamujú svoju správu implicitným spôsobom. Jedným zo spôsobov nepriamej reči je sarkazmus (alebo inak slovná irónia). Sarkazmus je bežný v užívateľsky tvorených obsahoch, ako napr. blogy alebo recenzie. Zo svojej podstaty dvojjazyčný charakter sarkazmu komplikuje jeho rozpoznávanie. Niekedy je aj pre ľudí zložitý rozoznať, či ide skutočne o sarkazmus alebo nie.

V dnešnej dobe je internet silný nástroj na komunikáciu medzi ľuďmi, alebo získania názoru pri kúpe nejakého produktu. Analýza týchto názorov [2] je často nepresná kvôli používaniu sarkazmu recenzentov a teda systém dolovania názorov môže nesprávne rozhodnúť o výslednej polarite príspevku.

## II. ROZPOZNÁVANIE SARKAZMU

Sarkazmus je dobre študovaný jav v lingvistiky, psychológii a v kognitívnych vedách, ale v textoch je jeho detekcia považovaná za ťažký problém a bol riešený len v niekoľkých štúdiách. V kontexte hovorených dialógov, automatická detekcia sarkazmu sa spolieha predovšetkým na rečovo súvisiace podnety, ako je smiech a prozódia (náuka o rytmickej, jazykovej a zvukovej stránke viet) [3].

Jedna z definícií sarkazmu je: schopnosť povedať alebo napísať niečo opačne od pôvodnej myšlienky alebo rozprávať spôsobom, aby sme niekoho donútili cítiť sa hlúpo alebo naznačiť niekomu, že sme nahnevaní. Okrem psychologického a vedeckého záujmu je automatické zisťovanie sarkazmu veľmi zaujímavé aj z komerčného hľadiska. Štúdie preferencií používateľov naznačujú, že hodnotenia, ktoré obsahujú sarkazmus považuje časť používateľov za menej nápomocné, aj keď na druhej strane, niektorí preferujú čítanie sarkastických recenzií [3]. Veľa štúdií o sarkazme hovorí, že sarkazmus vychádza z medzery medzi očakávaným vyjadrením a skutočným zveličením vyjadrenia a preháňaním, ako je to namodelované v teóriách zmienky a narážok. Implementoval sa základ navrhnutý zachytiť zmienku sarkazmu a pokúsiť sa splniť definíciu "hovoriť opak toho čo si myslíš, aby si niekomu spôsobil to, aby sa cítil hlúpo, v prejave tvojho rozhorčenia".

## III. IMPLEMENTÁCIA

Aby identifikácia sarkazmu bola dostatočne presná, bolo potrebné vytvoriť dostatočne veľkú trénovaciu množinu. Trénovacia množina obsahovala 600 recenzií a príspevkov z Amazonu, ktoré sa týkali rôznych produktov komentovaných v mikrobloggerovej sociálnej sieti Twitter. Tieto príspevky boli manuálne získané. Keďže pochopenie sarkazmu nie je väčšinou možné



z jedného slova vo vete, ale z celého obsahu a kontextu vety. Preto bolo potrebné vytvoriť vzory, ktoré charakterizujú sarkazmus vo vete. Na začiatku je dôležité povedať, že vytvorený klasifikátor pracuje iba s anglickými vetami, pretože angličtina je svetový jazyk, oveľa viac používaný ako slovenský. Niektoré vety boli trochu nezrozumiteľné pri používaní slangových a hovorových slov, z tohto dôvodu bolo ešte trochu ťažšie pochopiť charakter sarkazmu v anglických vetách. Teda pri identifikovaní sarkastických viet bolo nutné sa hlbšie vnoriť do významu vety. Taktiež pri identifikácii viet sa bral do úvahy takzvaný „star-rating“ každého hodnotiaceho používateľa, ktorý je povinný poskytnúť toto hodnotenie, aby sa dal identifikovať nespokojný používateľ.

Môžeme povedať, že veľkú časť sarkastických viet charakterizujú tieto spoločné črty:

- Počet výkričníkov vo vete
- Počet otáznikov vo vete
- Počet úvodzoviek vo vete
- Počet slov s veľkými písmenami vo vete

Vzory boli napísané do textového súboru vo forme regulárnych výrazov (regexov), ktoré boli podľa uváženia najlepším nástrojom na nájdenie vzoru v zadanom texte. V tomto príspevku uvažujeme 5 stupňov sarkazmu (1-5), pričom 1 znamená úplnú absenciu sarkazmu, 5 silný sarkazmus. Každý riadok v slovníku sa skladá z daného vzoru a k nemu je prislúchajúci stupeň sarkazmu. Stupne sarkazmu sa zadávali manuálne z danej vety s uvážením „star-ratingu“.

#### IV. EXPERIMENTY

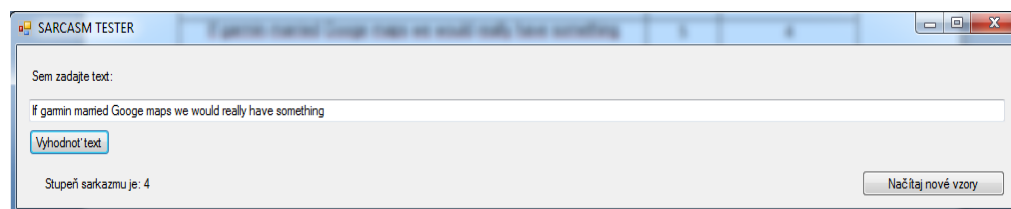
Experimenty boli vykonané na dátach o veľkosti 362 viet. Experiment bol rozdelený na dve časti. Prvý experiment bol vykonaný na vzorke o 182 viet získaných z Amazonu, druhý experiment bol vykonaný na vzorke o 180 viet získaných z Twittra.

##### A. Experiment pre dáta z Amazonu

V nasledujúcej tabuľke je ukázaný experiment na príkladoch viet z Amazonu a identifikovanie expertom a našim klasifikátorom (viď Obr.1):

Tabuľka 1  
Príklady sarkastických viet z Amazonu a ich ohodnotenie expertom a našim klasifikátorom

Vety	Expert	klasifikátor
By reading this book I now recognise what a brilliant writer John Steinbeck is.	1	1
I think they come with a free filter that omits all low range sounds so if you like listening to music with absolutely no Bass THESE HEADPHONES ARE FOR YOU.	5	3
Good luck guessing where the GPS wants you to go	5	5
I could have bought 3 to 4 pairs of the Sonys for what these cost	3	1
I'm almost never motivated to write a review but my set crapped out a few weeks after the warranty. Yay!	5	5
I'm smart and I'm diligent and these do not work for me.	4	2



Obr. 1 Príklad vyhodnotenia sarkastickej vety pre Amazon dáta

Následne diskretizáciou prevedieme Tabuľku1 do podoby ilustrovanej v Tabuľke 2. Pri diskretizácii uvažujeme A – sarkasticky pozitívne vety, ktoré majú stupne 3,4,5. N – sarkasticky negatívne vety majú stupne 1,2. Ak sú všetky vety takto pripravené, je možné zostaviť kontingenčnú tabuľku na vypočítanie presnosti a návratnosti podľa zdroja [1].

Tabuľka 2  
 Rozdelenie viet na sarkasticky pozitivne a sarkasticky negativne pre Amazon data

Vety	Expert	Klasifikator
By reading this book I now recognise what a brilliant writer John Steinbeck is.	N	N
I think they come with a free filter that omits all low range sounds so if you like listening to music with absolutely no Bass THESE HEADPHONES ARE FOR YOU.	A	A
Good luck guessing where the GPS wants you to go	A	A
I could have bought 3 to 4 pairs of the Sonys for what these cost	A	N
I'm almost never motivated to write a review but my set crapped out a few weeks after the warranty. Yay!	A	A
I'm smart and I'm diligent and these do not work for me.	A	N

Tabuľka 3  
 Kontigenčná tabuľka pre Amazon data

		Expert	
		A	N
Implementácia	A	67	15
	N	17	83

Pre Amazon data je výsledná návratnosť  $R = 0,797$ , presnosť  $\Pi = 0,817$  a F funkcia má hodnotu  $F_1 = 0,806$ .

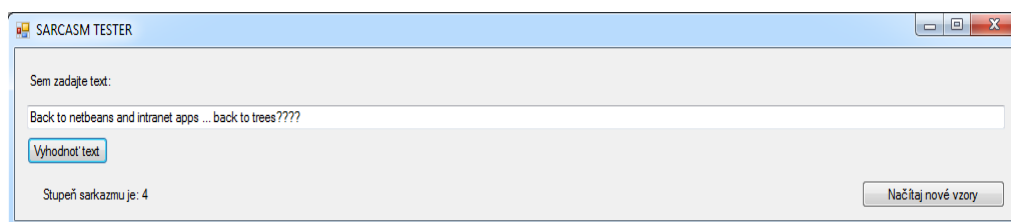
*B. Experiment pre data z Twittra*

V nasledovnej tabuľke je ukázaný experiment na príkladoch viet z Twittra a identifikovanie viet expertom a našim klasifikátorom (viď Obr.2):

Tabuľka 4  
 Príklady sarkastických viet z Twittra a ich ohodnotenie expertom a navrhnutým klasifikátorom.

Vety	Expert	Klasifikator
Have I mentioned I love my job? Even Leon the Chameleon's looking very green this morning. Still digesting his cockroaches, I guess.	5	3
Even if I understand and use a lot tags in photos, I still don't see the use of Gravityapps' Tags. Anybody is using it ?	2	3
School Board meeting 7 PM tonight! Shoreline Center.	1	1
I have a laaaame pick up line....Did it hurt? when you fell from heaven??? hee hee!	3	4
Bobbing around in Mahone Bay. Guys trying to save a sinking speed boat :)	2	1
Traded in my useless NEW phone today for another one. Technology is wonderful when it works	4	5

Následne diskretizáciou prevedieme Tabuľku4. do podoby.ilustrovanej Tabuľkou 5.



Obr. 2 Príklad vyhodnotenia sarkastickej vety pre Twitter data

Tabuľka 5  
 Rozdelenie viet na sarkasticky pozitivne a sarkasticky negativne pre Twitter dáta

Vety	Expert	klasifikátor
Have I mentioned I love my job? Even Leon the Chameleon's looking very green this morning. Still digesting his cockroaches, I guess.	A	A
Even if I understand and use a lot tags in photos, I still don't see the use of Gravityapps' Tags. Anybody is using it ?	N	A
School Board meeting 7 PM tonight! Shoreline Center.	N	N
I have a laaaame pick up line....Did it hurt? when you fell from heaven??? hee hee!	A	A
Bobbing around in Mahone Bay. Guys trying to save a sinking speed boat :)	N	N
Traded in my useless NEW phone today for another one. Technology is wonderful when it works	A	A

Tabuľka 6  
 Kontigenčná tabuľka pre Twitter dáta

		Expert	
		A	N
Implementácia	A	81	21
	N	11	67

Pre Twitter dáta je výsledná návratnosť  $R = 0,88$ , presnosť  $\Pi = 0,794$  a F funkcia má hodnotu  $F_1 = 0,832$ .

#### V. ZÁVER

Úlohou tohto príspevku bolo prezentovať návrh systému pre automatické rozpoznávanie sarkazmu v konverzačnom obsahu. Bol navrhnutý systém, ktorý analyzoval zadaný vstupný text v anglickom jazyku a jeho výsledkom bolo určenie stupňa sarkazmu danej vety. Náročné a zdĺhavé bolo vytvorenie vzorov z jednotlivých viet a ich ohodnotenie stupňom sarkazmu. Pri vytváraní vzorov do slovníka bolo potrebné mať vedomosti v oblasti regulárnych výrazov.

Pri testovaní programu, ktorý vyhodnotil spolu 362 viet, z toho 182 viet bolo z Amazonu, kde bola dosiahnutá presnosť 81% a 180 viet z Twittra, kde bola dosiahnutá presnosť 79%. Tieto presnosti sú prekvapivo vysoké vzhľadom na náročnosť identifikovania sarkazmu. Napriek tomu by bolo žiaduce dosiahnuť v budúcnosti vyššiu presnosť. Chyby mohli byť spôsobené podobnou trénovacou vzorkou, keďže bol použitý rovnaký zdroj, alebo nedostatočne veľkou testovacou vzorkou, prípadne nesprávnym vytvorením vzoru a určením jeho stupňa.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### ZOZNAM POUŽITEJ LITERATÚRY

- [1] K. Machova, "Strojové učenie v systémoch spracovania informácií." ELFA s.r.o., 2009, Košice, 85s, ISBN 978-80-8086-130-8.
- [2] M. Russell, "With sentiment analysis, context always matters." <http://radar.oreilly.com/2011/03/sentiment-analysis-context.html>
- [3] O. Tsur, "Rappoport, A.: ICWSM – A Great Catchy Name : Semi-Supervised Recognition of Sarcastic Sentences" in Online Product Reviews, Dostupné na internete: <http://staff.science.uva.nl/~otsur/papers/sarcasmAmazonICWSM10.pdf>

# Porovnanie metód transformácie kardinalitných ohraničení na SAT problémy

<sup>1</sup>Veronika FEKETEOVÁ, <sup>2</sup>Marián MACH

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>veronika.feketeova@student.tuke.sk, <sup>2</sup>marian.mach@tuke.sk

**Abstrakt** — Cieľom článku je problematika transformácie ohraničení do SAT podoby. Bolo vytvorené testovacie prostredie pre transformáciu ohraničení. V rámci tohto prostredia bola vytvorená sada programov pre transformáciu ohraničení na maximálny počet vybraných prvkov (označovaných ako kardinalitné ohraničenia) do SAT podoby. Tieto ohraničenia boli transformované na základe štyroch vybraných a následne implementovaných transformačných metódach. Týmito transformačnými metódami sú metóda sekvenčného kódovania, metóda binomického kódovania, metóda commander kódovania a poslednou metódou je binárne kódovanie. V rámci experimentovania s vytvoreným testovacím prostredím bolo realizované vzájomné porovnanie kódovacích metód.

**KLúčové slová** — CNF formát, SAT problém, splniteľnosť, transformácia ohraničení

## I. ÚVOD

SAT problém je takzvaný problém splniteľnosti Booleovských formúl [1]. Ako vstup je daná takáto formula (kódujúca problém ktorý je potrebné riešiť) a my sa pýtame, či je táto formula splniteľná. To znamená, že sa hľadá také priradenie logických hodnôt jednotlivým premenným, aby výsledná formula bola splnená.

Pretože pre riešenie SAT úloh sú k dispozícii moderné výkonné „SAT solvery“, je možné toto využiť aj pri riešení iných úloh – riešená úloha sa transformuje do podoby SAT problému, ten sa vyrieši a riešenie je spätne transformované na riešenie pôvodnej úlohy. Jedným z typov úloh, kde túto cestu možno využiť, sú úlohy s ohraničeniami. V tomto prípade je potrebné stanoviť, ako sa rôzne typy ohraničení majú transformovať na SAT úlohu. Často využívanými sú tzv. kardinalitné ohraničenia, kde je potrebné vybrať aspoň/práve/najviac určitý počet prvkov z množiny prvkov ktoré sú k dispozícii. Dôraz v tomto článku je na „najviac“ ohraničení a jeho transformácii do podoby klauzúl CNF reprezentácie logickej formuly.

Článok je organizovaný nasledovne. Po stručnom predstavení vybraných ohraničení (jedného základného a troch alternatívnych) je predstavené vytvorené testovacie prostredie zamerané na dva typy testovacích úloh. Následne je uvedený popis experimentov spolu s dosiahnutými výsledkami a ich zhodnotením.

## II. VYBRANÉ TRANSFORMAČNÉ METÓDY

### A. Metóda binomického kódovania

Toto je základná metóda kódovania. Vychádza z predpokladu, že ak je možné vybrať najviac  $k$  prvkov (teda najviac  $k$  premenných z nejakej skupiny logických premenných môže byť pravdivých), tak nie je možné vybrať  $k+1$  prvkov, pretože v prípade takéhoto výberu by dané ohraničenie bolo porušené. Aby sa zabránilo výberu nejakej  $(k+1)$ -tice prvkov  $X_1$  až  $X_{k+1}$ , stačí vygenerovať klauzulu tvaru „ $\neg X_1 \vee \neg X_2 \vee \dots \vee \neg X_k \vee \neg X_{k+1}$ “ (disjunkcia negácií premenných). A toto musí platiť pre každú  $(k+1)$ -ticu.

Toto hovorí, že sú vytvárané jednotlivé kombinácie všetkých premenných zo vstupu a pre každú z nich je generovaná klauzula uvedeného tvaru (všetky premenné v kombinácii sú zároveň negované a medzi nimi je disjunkcia). Tvorba jednotlivých kombinácií závisí od

hodnoty  $k$ . Ak je  $k=1$  tak sú vytvárame všetky možné dvojice, ak  $k=2$  tak sú vytvárame všetky možné trojice a takto to ide ďalej [2].

Nevýhodou tohto základného spôsobu kódovania je to, že počet generovaných klauzúl kombinačne rýchlo rastie s rastúcou hodnotou  $k$  a počtom premenných, ktoré sú k dispozícii. Preto sa hľadajú alternatívne metódy kódovania, ktoré vedú na menší počet generovaných klauzúl aj za cenu pridávania pomocných premenných a tým pádom zväčšovania priestoru, v ktorom je potrebné hľadať riešenie úlohy.

#### B. Metóda kódovania sekvenčným čítačom

Metóda vychádza z predstavy obvodu sekvenčného čítača, pozostávajúceho zo skupiny registrov, ktoré v sebe obsahujú AND a OR hradlá [3]. Jednotlivé vstupy a výstupy registrov (pomocné premenné) sú prepojené tak, že  $j$ -ty výstup  $i$ -teho registra nadobúda pravdivú hodnotu vtedy, ak aspoň  $j$  z prvých  $i$  premenných je pravdivých. Chovanie obvodu je vyjadrené pomocou piatich typov klauzúl, ktoré reprezentujú transformačný predpis.

#### C. Metóda binárneho kódovania

Metóda zavádza skupinu pomocných premenných, pričom každej pôvodnej premennej zodpovedá určitá jedinečná kombinácia hodnôt týchto pomocných premenných [2]. Toto priradenie je vyjadrené klauzulami, ktoré reprezentujú transformačný predpis.

#### D. Metóda kódovania commander

Metóda je založená na princípe „rozdeľuj a panuj“ – delí pôvodné premenné do skupín, pričom každej skupine priradí sadu pomocných premenných reprezentujúcich danú skupinu [4]. Takýmto spôsobom ohraničenie nad pôvodnými premennými je nahradené ohraničením nad menšou skupinou pomocných premenných a lokálnymi ohraničeniami nad jednotlivými skupinami. Zavedenie pomocných premenných a jednotlivé náhradné ohraničenia vytvárajú transformačný predpis pre ohraničenie nad pôvodnými premennými.

### III. POPIS TESTOVACIEHO PROSTREDIA

Ako testovacie prostredie pre experimentovanie s jednotlivými kódovacími metódami pri riešení úloh boli zvolené dva problémy, ktorými sú „holuby“ [4] a „sociálny golfista“ [5].

Tabuľka 1  
Testovací problém „Holuby“

Cieľom je rozmiestniť zadaný počet holubov do zadaného počtu otvorov tak, aby nebola prekročená kapacita ani jedného z otvorov.	
<i>Parametre</i>	
$p$	počet holubov
$h$	počet otvorov
$k$	kapacita každého otvoru
<i>Ohraničenia</i>	
1	Každý holub musí byť umiestnený aspoň do jedného otvoru (výber aspoň jedného prvku z celkového počtu $h$ prvkov)
2	Každý holub musí byť umiestnený najviac do jedného otvoru (výber najviac jedného prvku z celkového počtu $h$ prvkov)
3	V každom otvore môže byť umiestnených najviac $k$ holubov (výber najviac $k$ prvkov z celkového počtu $p$ prvkov)

Testovací problém „Holuby“ je založený na umiestňovaní holubov do dier (Tab. 1), pričom je potrebné kódovať tri typy ohraničení. V prípade ak platí, že počet holubov je väčší alebo rovný ako spoločná kapacita otvorov ( $p \geq k \cdot h$ ), tak nie je potrebné explicitne kódovať ohraničenie 2, pretože jeho platnosť vyplýva z platnosti ostatných ohraničení. Ak počet holubov prevyšuje kapacitu otvorov ( $p > k \cdot h$ ), tak úloha je nesplniteľná, v opačnom prípade je splniteľná (a pomerne ľahká kvôli existencii veľkého počtu navzájom symetrických riešení).

Testovací problém „Sociálny golfista“ je založený na rozdelení hráčov v golfovom klube tak, aby si každý hráč zahral s čo najväčším počtom ostatných členov klubu (Tab. 2). Zložitost' problému závisí od počtu hracích kôl – problém je triviálny pre jedno hracie kolo, so

stúpajúcim počtom hracích kôl je stále zložitejší, až sa pri určitom počte kôl stáva nespľniteľným.

Tabuľka 2  
 Testovací problém „Sociálny golfista“

Cieľom je zoskupovať zadaný počet hráčov v každom hracom kole do hracích skupín tak, aby žiadny hráč nehral počas zadaného počtu kôl so žiadnym iným hráčom viac ako raz.	
<i>Parametre</i>	
$p$	veľkosť skupiny (počet miest v skupine)
$g$	počet skupín
$w$	počet hracích kôl
<i>Ohraničenia</i>	
1	Každý hráč je v každom hracom kole zaradený aspoň do jednej skupiny (výber aspoň jedného prvku z celkového počtu $g$ prvkov)
2	Každý hráč je v každom hracom kole zaradený najviac do jednej skupiny (výber najviac jedného prvku z celkového počtu $g$ prvkov)
3	V každej skupine je v každom hracom kole zaradených aspoň $p$ hráčov (výber aspoň $p$ prvkov z celkového počtu $p * g$ prvkov)
4	V každej skupine je v každom hracom kole zaradených najviac $p$ hráčov (výber najviac $p$ prvkov z celkového počtu $p * g$ prvkov)
5	Každá dvojica hráčov hrá spolu najviac raz (výber najviac jedného prvku z celkového počtu $g * w$ prvkov)

#### IV. REALIZOVANÉ EXPERIMENTY

Cieľom tohto experimentu bolo porovnať na daných úlohách všetky uvažované kódovacie metódy navzájom. Danými kódovacími metódami boli testovacie úlohy zakódované do tvaru SAT problému reprezentovaného v Dimacs CNF formáte a následne riešené vybraným SAT solverom [6]. Meranou charakteristikou bol čas, ktorý solver potreboval na vyriešenie úlohy alebo rozpoznanie, že úloha nie je splniteľná.

Pre overovanie úspešnosti jednotlivých spôsobov kódovania pre prípad nespľniteľných úloh bol použitý testovací problém „Holuby“ v konfigurácii, keď celkový počet holubov bol väčší ako celková kapacita otvorov. Keďže bol použitý solver založený na systematickom prehľadávaní priestoru, bol nútený prehľadať celý priestor prehľadávania pred tým ako mohol indikovať nespľniteľnosť riešenej úlohy.

Tabuľka 3  
 Priemerné časy riešenia nespľniteľných úloh (problém “Holuby”)

$p$	$h$	$k$	Sekvenčne	Binomicky	Commander	Binárne
7	6	1	0.06	0.21	0.18	0.14
8	7	1	0.07	0.27	0.27	0.18
9	8	1	0.27	0.79	0.76	0.34
10	9	1	1.82	2.19	2.22	2.38
11	10	1	23.43	30.97	29.73	11.72
12	11	1	116.48	85.84	85.08	52.95
13	12	1	2774.66	456.82	447.76	1879.57
14	13	1	-	1300.89	1236.16	-
7	2	3	0.20	0.14	0.15	0.17
10	3	3	0.14	0.20	0.19	0.20
13	4	3	0.47	510.54	512.11	0.28
16	5	3	11.98	-	-	1.31
19	6	3	698.98	-	-	40.50
22	7	3	-	-	-	1819.88

Pomlčka znamená, že v danom prípade výsledok nebol dosiahnutý do limitu 3600 sekúnd.

Ukážka dosiahnutých výsledkov je v Tab. 3 (ďalšie hodnoty je možné nájsť v [7]). Keďže

nie sú dôležité absolútne hodnoty jednotlivých časov ale porovnania časov navzájom (s kvantifikáciou „približne rovnaké“, „rýchlejší“, „oveľa rýchlejší“), tak nie je potrebné uvádzať technické špecifikácie použitého stroja. Na základe tohto testovania bolo zistené, že pri riešení úlohy, ak riešenie neexistuje, tak v prípade  $k=1$  (ohraničenie na jeden prvok) je najvhodnejšia metóda kódovania commander kódovanie alebo binomické kódovanie. Ale ak  $k>1$  (ohraničenie na viac prvkov), tak najvhodnejšou metódou je pre vyriešenie danej úlohy binárne kódovanie nasledované sekvenčným kódovaním.

Pre overovanie úspešnosti jednotlivých spôsobov kódovania pre prípad splniteľných úloh bol použitý testovací problém „Sociálny golfista“ v konfigurácii piatich skupín po troch hráčoch, keď počet hracích kôl bol volený tak, aby vznikla konfigurácia o ktorej je známe že je splniteľná. V tomto prípade použitý solver bol nútený prehľadať iba časť celého priestoru prehľadávania pred tým ako mohol indikovať splniteľnosť riešenej úlohy.

Tabuľka 4  
Priemerné časy riešenia splniteľných úloh (problém “Sociálny golfista”)

$p$	$g$	$w$	Sekvenčne	Binomicky	Commander	Binárne
3	5	1	0.29	0.67	0.31	0.55
3	5	2	0.50	1.16	0.50	3.54
3	5	3	0.71	1.37	0.75	4.99
3	5	4	1.06	1.71	1.08	5.82
3	5	5	1.30	2.00	1.24	10.43
3	5	6	-	3.48	1.59	-
3	5	7	-	27.16	318.85	-
3	5	8	-	-	-	-

Pomlčka znamená, že v danom prípade výsledok nebol dosiahnutý do limitu 3600 sekúnd.

Ukážka dosiahnutých výsledkov je v Tab. 4. Na základe tohto testovania bolo zistené, že pri riešení úlohy, ak počet hracích kôl je malý, sú jednotlivé metódy viac menej rovnocenné (s výraznejším zaostávaním binárneho kódovania). Ale ak počet kôl narastá, tak sekvenčné kódovanie stráca kontakt a najvhodnejšou metódou pre vyriešenie danej úlohy sa stáva binomické kódovanie nasledované kódovaním commander.

Osobitným prípadom bola konfigurácia s ôsmimi hracími kolami. Pre túto konfiguráciu neexistuje vhodné riešenie (úloha je nesplniteľná pretože ak by každý hráč mal počas týchto kôl hrať s niekým iným, tak by počet hráčov musel byť aspoň 17 a nie iba 15). Toto však neumožnilo vo vymedzenom čase 3600 sekúnd odhaliť ani jedno z použitých kódovaní.

## V. ZÁVER

V práci bolo na dvoch testovacích úlohách overené, že voľba vhodnej metódy pre transformáciu kardinalitných ohraničení na SAT klauzuly nie je jednoduchou, pretože neexistuje taká metóda, ktorá by bola vhodná na každý typ problému. Zároveň práca predkladá závery, týkajúce sa použitia štyroch testovaných transformačných metód a delí ich do skupín – na metódy vhodné pre nesplniteľné úlohy, splniteľné úlohy menšej zložitosti (s menším počtom klauzúl) a pre splniteľné úlohy väčšej zložitosti (s väčším počtom klauzúl).

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] A. Biere et al, *Hanbook of Satisfiability*, Fontiers in Artificial Intelligence and Applications, vol 185, IOS Press, Amsterdam, 2009.
- [2] A.M. Frisch and P. A. Giannaros, “SAT Encodings of the At-Most-k Constraint: Some old, some new, some fast, some slow”, in *Proc. of the 9th Int. Workshop on Constraint Modelling and Reformulation (ModRef2010)*, 2010.
- [3] C. Sinz, “Towards an optimal CNF encoding of Boolean cardinality constraints”, in *Proc. of the 11th Int. Conf. on Principles and Practice of Constraint Programming (CP 2005)*, Sitges, Spain, October 2005, 827-831.

- [4] W. Kieber and G. Kwon, "Efficient CNF Encoding for Selecting 1 from N Objects", in *Fourth Workshop on Constraints in Formal Verification*, Bremen, Germany, July 2007.
- [5] I. P. Gent and I. Lynce, "A SAT Encoding for the Social Golfer Problem", in *IJCAI'05 Workshop on Modelling and Solving Problems with Constraints*, Edinburgh, Scotland, July 2005.
- [6] Cryptominisat 2 SAT solver, <http://www.msoos.org/cryptominisat2/> [10.7.2014]
- [7] V. Feketeová, "Transformácia úloh s ohraničeniami na SAT problem", Bakalárska práca, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, 2014, 55 strán.



# Evolutionary agent model creation

<sup>1</sup>Viliam LITAVEC, <sup>2</sup>Marián MACH

<sup>1,2</sup> Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>viliam.litavec@student.tuke.sk, <sup>2</sup>marian.mach@tuke.sk

**Abstract** — The paper examines the issue of creating agent models using evolutionary algorithms. A common approach is to represent each agent model as a set of rules. To produce such rules from behavior of a modeled agent, an evolutionary algorithm for learning rules was employed – specifically the SIA01 algorithm was used. We designed a process cycle as we proceeded to create agent models. We created a software game simulator, which enables to generate data for learning procedure. From the data obtained, we learned the model using the algorithm SIA01. Subsequently, the learned models could be utilized within the simulator with the possibility to compare the performance of the learned models. Finally, we carried out experiments where we observed the development of models during iterations as they change, and tested how they perform when they are applied.

**Keywords** — agent model, evolutionary algorithm, rule learning, SIA01

## I. INTRODUCTION

All live beings as well as all artifacts produced by people can be represented as agents. If we try to replace some autonomous entities and create agents as replacements for them, they must act like original ones. Thus, it is necessary to create agent behavior models. We consider an agent model like a set of rules which serve as a tool to decide how to act in a specific situation. To create such set of rules we need to use rule learning algorithms. For our needs we have selected evolutionary algorithms to be used. With these algorithms at disposal we observed how the agent behavior models can be created, applied and whether they can be improved, how input data affect the models and what is their performance and what affects them.

## II. DESCRIPTION OF EVOLUTIONARY ALGORITHM SIA01

To solve a given problem, we have chosen the algorithm SIA01. The algorithm is used for supervised learning first order logic rules using evolutionary search for the learning. It is a successor of the SIA algorithm [1] which uses more limited attribute representation. In contrast, the SIA01 algorithm [2] uses a high level representation for learning the rules of the first order logic and may deal with numeric data as well as background knowledge like hierarchies or tree structured values. These assumptions helped us in the selection.

The algorithm works with a representation of rules of the form: “description implies class” where the description part comprises predicates (together with constants and variables) describing some situation, and the class part describes the decision that should occurred in this situation. For the purpose of this paper it was necessary that the individual elements with which algorithm works are of fixed length. Elements will be composed of the chromosomes where predicates and their arguments are genes. Individual chromosomes represent values of attribute representation, which is templated at the beginning of the work.

### A. Covering Principle

This principle is the same as in the AQ algorithm (the principle is known as star methodology). On input there is a set of examples to work with. From this example set a sample is randomly selected. The system then tries to find the best rule from the sample while the rule is evaluated according to its fitness. The aim is to use this rule to cover the greatest number of examples. If there is any example covered by the rule, it is removed from the example set. The developed rule is stored in the output variable. This process is repeated until

the example set is not empty. After the completion of the process a set of rules covering the initial set of examples is available.

### B. SIA01 Search Principle

This principle starts the job with an initial population of the selected sample of the example set. The objective is to create the best population of rules. In this context the search is performed within a cycle where genetic operators are applied on all members of the population. Mutation operator in individual components will replace elements for other elements using uniform mutation. One-point crossover can be used in the role of the crossover operator.

Within the application of evolutionary operators on a population member three new members of the population are being created and temporarily stored. After the end of the evolution cycle, a new generation will be formed - new members can be added to the population. This addition can be performed under these conditions: if the population size is less than the desired size (for addition) or if fitness of a new member is better than fitness of the worst member of the population (for replacement). The condition for ending the evolutionary cycle is the inability to add a new member during some number of generations.

## III. TEST ENVIRONMENT

To experiment with agent model creation, an experimental software playground has been developed [3]. It has the following cyclic structure:

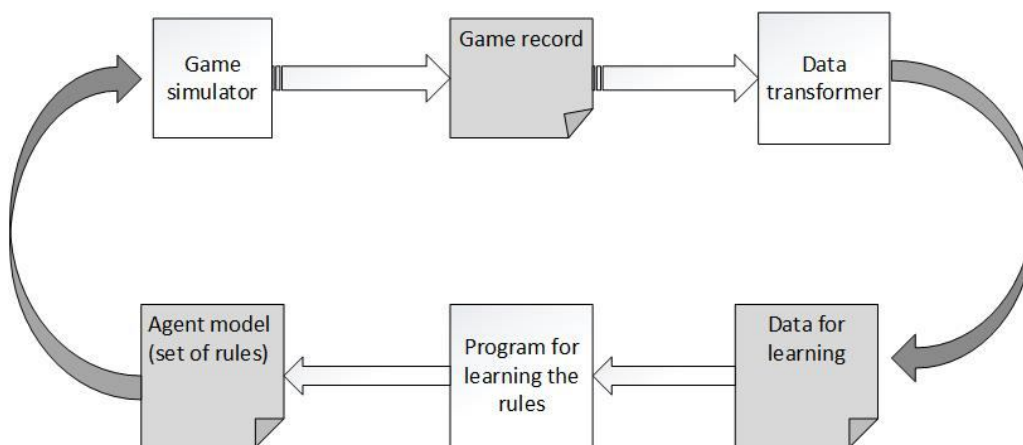


Fig. 1 Experimental software playground structure.

At the ‘beginning’ of the experiment cycle there is a program simulator for the game known as “Ludo” (or “Mensch ärgere dich nicht” in German). The program simulates a game of two players. The task of the program is to create a log describing how a particular game proceeded. It logs the play from the perspective of both players separately and thus creates two separate record files. These data must be changed to attribute representation before they can be used as input to the learning algorithm. To automate this process, a simple program is used to transform the data into the desired shape. Its task is to select and process the data from the game record and save them into a new file in the required form. The program will create a file with data that can be used to learn the rules. Subsequently, a learner (program implementing the SIA01 algorithm) finds a set of rules that cover input data. This resulting set of rules represents an agent model. To test whether a proper model was created, it is possible to insert the generated set of rules as a behavior model representing one player into the game simulator. Performing this cycle once is called iteration.

The game simulator enables to define the behavior of agents. It is possible to define a model for single player or for both players or for neither of them. If no model for a player is defined, then the behavior of the player is based on the default model wired in the program, representing a random strategy. To compare the performance of models, it is possible to challenge two agents in order to play a tournament. The tournament is the simulation of a sequence of several games where a score is calculated for each player. This allows observing the performance of various learned models in mutual confrontations.

Before attribute-based model learning can start, a proper representation of attributes must be determined in order to make a correct model (model cannot use an attribute which is not defined). Therefore, it was necessary to initially establish the most appropriate attributes that should be considered in the search process of rule learning. For example attributes can evaluate positions of individual tokens on the board, relationships between token positions, dice rolls, whether player has tokens in his yard, and so on. Based on these attribute definitions, record data are transformed into examples from which it is possible to learn the rules.

Searching for rules follows the procedure mentioned in the description of the algorithm. This creates a population with a maximum size of 50 members. Evolution ends when a given number of generations (specified by user) does not add any new member to the population. Subsequently, the best rule is selected and serves as the output of the evolutionary part. If all data are covered by rules, the output is a collection of rules that can be called a model of the agent's of behavior. The model can be used as a strategy for an agent in the game simulator to play in tournaments or to produce log records related to the model.

#### IV. FIRST EXPERIMENT

This experiment observes how model is being developed. Models are iterated through the loop and therefore it is possible to observe how that process evolves. For comparison, we created models from different initial strategies - four basic strategies were used (the first strategy raced one token only, the second strategy added further token into play when possible, the third strategy was offensive and the fourth was default random strategy). Each strategy was iterated three times. In the experiment we observed how rules had evolved in these iterations, whether iterating through models gives better results. In the evaluation of the experiment, we distinguish two types of rules - correct rules and inaccurate rules. Inaccurate rules are rules that are very specific or rules that decide on moving a token while taking into account conditions unrelated to the given token. Classification of rules was based on personal opinion and experience gained during the work.

Table 1  
Results of the first experiment

	Model	Rules	Correct	Inaccurate	% Correct	% Inaccurate	Rules difference	% Difference
1 Iteration	1	18	15	3	84 %	16 %		
	2	18	16	2	89 %	11 %		
	3	15	14	1	94 %	6 %		
	4	15	15	0	100 %	0 %		
3 Iterations	1	16	15	1	94 %	6 %	-2	+10 %
	2	13	13	0	100 %	0 %	-6	+11 %
	3	16	15	1	94 %	6 %	+1	0 %
	4	14	13	2	93 %	7 %	-1	-7 %

Based on the results in Tab. 1, it is possible to evaluate improvements provided by iterating the model through more iterations. The table presents the removal or reduction of the number of inaccurate rules resulting in increase of the accuracy of finding the right rules in the model in later iterations. But the results were very good already in the first iteration so the improvement is not so significant. When comparing different models within different iterations, learned rules are very similar. Since the rules learned are generalized, there is not a large space for rules to be evolved and changed within higher number of additional iterations.

#### V. SECOND EXPERIMENT

In another experiment, we compared the performance of different models (created in different ways) against each other. To make the comparison, the simulator was used for performing tournaments which challenged pairs of models. Two types of models were tested: models created within a single iteration but using different initial strategies as well as models created after 1st iteration versus models from 3rd iteration. In addition, influence of the used learning algorithm was investigated as well. To compare differences in the algorithms, we

used the models learned using SIA01 (evolutionary algorithm) and CN2 (machine learning algorithm). These models played against each other in tournaments consisting of 4 x 10 games.

*A. Models based on one iteration only*

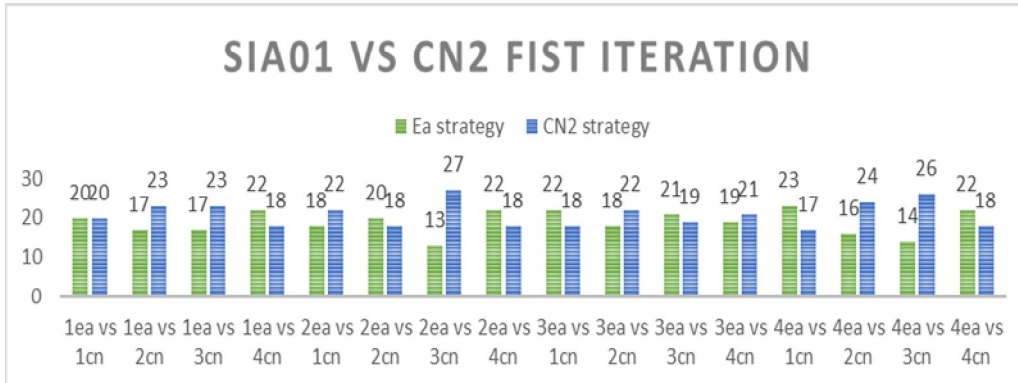


Fig. 2 Confrontation of models made with SIA01 and CN2.

Fig 2 presents results of all tournaments. The left bar in each pair of bars (a tournament) represents score of the model learned by evolutionary algorithm. The right bar of each tournament represents the score obtained by the model learned using CN2 algorithm. The description below each tournament indicates which strategies the models, participating in the tournament, were based on. From the results it can be seen that after the first iteration majority of battles were won by models learned using the CN2 algorithm. In addition, there was no opportunity to win all the tournaments using one strategy against the other strategies. But certain strategies combined with CN2 were able to win by bigger difference than strategies combined with evolutionary algorithm.

*B. Models based on different numbers of iterations*

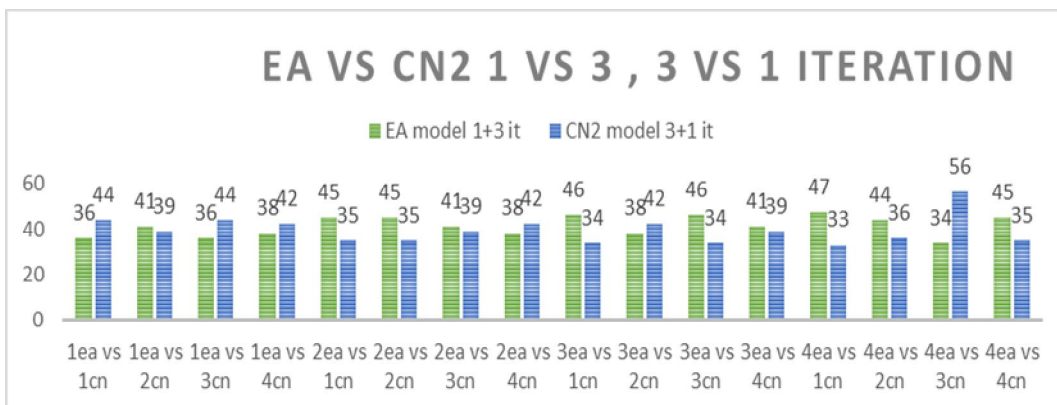


Fig. 3 Confrontation of models (1 vs 3 iterations) made with SIA01 and CN2.

Fig 3 presents the results in the same way as the previous chart. The difference is that one pair of bars stands for two tournaments – in the first tournament the first model was from the 1st iteration and the second model from the 3rd iteration, while in the second tournament the situation was reverted, the first model originated from the 3rd iteration and the second model from the 1st iteration. From the results it can be seen an interesting change. The performance of the models learned using evolutionary algorithm improved significantly – now they won majority of duels. Although still there is no strategy to be a clear winner better than the other strategies.

To summarize the results of the experiments, it is possible to claim that there are pairs of models based on different strategies, relative performance of which remains the same regardless of the used number of iterations. When considering the result for higher number of iterations, it was found that the iterated learning scenario showed better results for models learned using evolutionary algorithm – models were able to improve their performance and to

win over the opponent in the tournament in contrast to using only one iteration when they were not strong enough to defeat the opponent.

## VI. CONCLUSION

It was possible to successfully create agent models and subsequently to apply them within the game simulator. Based on the experiments carried out we have reached the following conclusions. When creating models it is appropriate to seek for rules with 1-3 conditions. These rules are generalized enough to cover a large number of situations and therefore they are very useful. In the process of iterating it is possible to reach improvement of created models. Using different algorithms for learning the rules was observed beneficial. When iterating using evolutionary algorithm, it was possible to remove inaccurate rules. Usage of CN2 algorithm provided more generalized rules but with low possibility to reduce the number of inaccurate rules. Regarding the performance of the obtained models interesting results were achieved as well. Iterating models through several iterations enabled to achieve improved results in the performance of the generated models – the increase of the number of iterations was more beneficial for models learned by evolutionary algorithm.

## ACKNOWLEDGMENT

This work is the result of the project implementation: Development of the Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120030) supported by the Research & Development Operational Program funded by the ERDF.

## REFERENCES

- [1] G. Venturini, "SIA: a supervised inductive algorithm with genetic search for learning attribute based concepts", in *Proc. of the CMP93*, 1993, pp. 280-296.
- [2] S. Augier, G. Venturini and Y. Kodratoff, "Learning first order logic rules with a genetic algorithm", in *Proc. of 1st Int. Conference on Knowledge Discovery & Data Mining*, AAAI Press, 1995, pp. 21-26.
- [3] V. Litavec, "Evolutionary creation of agent models", Master theses, Technical University of Kosice, 2014, 75 pages.

# Analýza mechanických vlastností LTCC substrátov

<sup>1</sup>Alena PIETRIKOVÁ, <sup>2</sup>Peter LUKÁCS

<sup>1,2</sup> Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika

<sup>1</sup>alena.pietrikova@tuke.sk, <sup>2</sup>peter.lukacs@tuke.sk

**Abstrakt** — Táto práca opisuje najčastejšie používané metódy analýzy mechanických vlastností troch typov LTCC (nízko teplotne vypaľovaná keramika) substrátov (DuPont 951, DuPont 9K7, MuRata LFC) založené na metóde trojbodového a štvorbodového ohybu a taktiež na DMA analýze (Dynamic Mechanical Analysis). Výsledky analýz ukazujú, že štvorbodový ohyb indikuje rozloženie pnutia v materiáli s vyššou presnosťou. Metóda trojbodového ohybu dosahuje vyššie hodnoty ohybovej pevnosti testovaných substrátov. Proces opakovaného výpalu LTCC substrátu DuPont 951 zvýšila hodnoty ohybovej pevnosti materiálu o 10 %. Zmena dĺžky trvania vypalovacieho profilu zo 4 na 26,5 hodín spôsobila zvýšenie ohybovej pevnosti materiálu o 100 %. Táto práca dokazuje, že MuRata má najlepšiu homogenitu a LTCC 9K7 má najväčšiu náchylnosť na krehko-lomový stav pri rozmedzí dolných nastavcov testovacieho systému 32 mm. Podmienky vypalovacieho procesu majú výrazný vplyv na zmršťovanie všetkých typov LTCC substrátov. DMA analýza LTCC substrátov v pružnej oblasti potvrdila výsledky získané trojbodovým a štvorbodovým ohybom.

**Kľúčové slová** — LTCC, trojbodový ohyb, štvorbodový ohyb, ohybová pevnosť, DMA analýza.

## I. ÚVOD

Zavedenie nízko-teplotne vypalovanej keramiky (LTCC) typu DuPont 9K7 a MuRata LFC v oblasti UWB radarových aplikácií prináša nové možnosti aplikácií z dôvodu výhodných dielektrických vlastností substrátov. LTCC technológia umožňuje výrobu rôznych komponentov pre oblasť mikrovln, ako sú RF filtre, antény, rezonátory, atď. Vývoj nových typov LTCC substrátov pre oblasť mikrovln je zameraný na vysoký faktor kvality, nízku relatívnu permitivitu substrátu a minimálne dielektrické straty. Dobré dielektrické vlastnosti LTCC substrátov ponúkajú možnosť miniaturizácie vysokofrekvenčných komponentov. Výhodou LTCC substrátov je nízky koeficient teplotnej rozťažnosti a malá drsnosť povrchu. Táto práca sa taktiež venuje vplyvu procesu výpalu na mechanické vlastnosti LTCC substrátov. Je všeobecne známe, že podmienky výpalu LTCC vyžadujú veľmi presné dodržanie teplotného profilu. Rozdielne mechanické vlastnosti viacvrstvových LTCC štruktúr sú zvyčajne zapríčinené odlišnými vypalovacími procesmi.

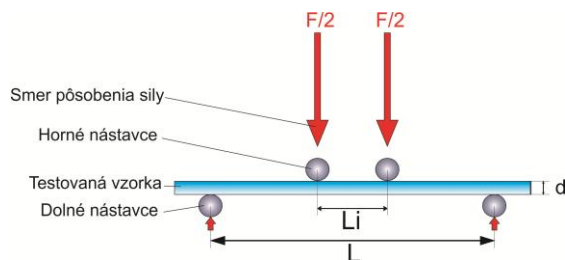
Hlavná myšlienka tejto práce vychádza z potreby pochopenia stability mechanických vlastností LTCC substrátov meraných trojbodovými a štvorbodovými metódami testovania mechanických vlastností materiálov. Ohybová pevnosť materiálu je ovplyvnená náchylnosťou na rýchly lom a pomalým rastom trhlin a mechanických porúch v materiáli. Motiváciou k tejto práci bola potreba analyzovať rôzne metódy merania mechanických vlastností materiálov z dôvodu možnosti predikcie stabilných vlastností LTCC substrátov v korelácii s hustotou a procesom kryštalizácie pri nízkych teplotách. Anizotropia vlastností má výrazný vplyv nie len na nestabilné mechanické vlastnosti LTCC substrátov, ale navyše aj na faktor kvality elektronických štruktúr v oblasti vysokých frekvencií.

## II. METÓDY MERANIA MECHANICKÝCH VLASTNOSTÍ LTCC SUBSTRÁTOV

Pre účely analýz mechanických vlastností rôznych typov LTCC substrátov boli použité trojbodové a štvorbodové metódy testovania ako aj DMA analýza. Trojbodový a štvorbodový ohyb sa zvyčajne používa na analýzu tendencie na krehko-lomový stav keramických materiálov. Bola taktiež použitá DMA analýza ako nový prostriedok na porovnanie nameraných výsledkov.

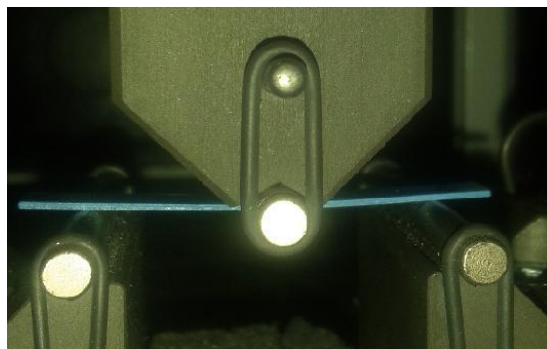
### A. Ohybové skúšky

Ohybové skúšky sa používajú na stanovenie mechanických vlastností krehkých materiálov, ako sú aj LTCC substráty. Trojbodový a štvorbodový ohyb slúži na zistenie ohybovej pevnosti materiálu a modulu pružnosti [3]. Princiálna schéma štvorbodového ohybu je znázornená na Obr. 1.



Obr. 1 Princiálna schéma štvorbodového ohybu.

Metóda štvorbodového ohybu indikuje rozloženie mechanického pnutia v materiáli s vyššou presnosťou a poukazuje na náchylnosť na krehko-lomový stav. Obr. 2 znázorňuje reálne meranie mechanických vlastností LTCC substrátu DuPont 951 metódou trojbodového ohybu.



Obr. 2 Detail merania mechanických vlastností LTCC substrátu DuPont 951 metódou trojbodového ohybu.

Ohybová pevnosť  $\sigma_{fs}$  sa vypočítava na základe sily  $F$ , ktorá v prípade trojbodového ohybu predstavuje vzťah 1. Pre metódu štvorbodového ohybu platí vzťah 2.

$$\sigma_{fs} = \frac{3FL}{2bd^2} \quad (1)$$

$$\sigma_{fs} = \frac{3F(L-L_i)}{2bd^2} \quad (2)$$

kde  $F$  je sila v bode lomu vzorky [N],

$L$  je rozmedzie medzi nástavcami [mm],

$b$  je šírka vzorky [mm],

$d$  je hrúbka vzorky [mm].

V prípade nami realizovaných meraní boli podmienky testovania rovnaké pre trojbodový aj štvorbodový ohyb. Rýchlosť pohybu horných nástavcov bola 1 mm/min a rozmedzie dolných nástavcov bolo 32 mm. Naše merania boli realizované na zariadení Testometric M250-2.5CT.

### B. DMA Analýza

Dynamická mechanická analýza (DMA) môže byť zafinovaná ako pôsobenie špecifikovanej sily na vzorku a analýza reakcie materiálu na túto silu. Veľmi dôležité je, že táto atypická testovacia metóda ponúka možnosť precíznej analýzy mechanických vlastností materiálu v pružnej oblasti. Sila pôsobiaca na vzorku sa označuje ako záťaž a vzorka vystavená tejto záťaži vykazuje deformáciu [5]. Obr. 3 znázorňuje upevnenie vzorky do testovacieho

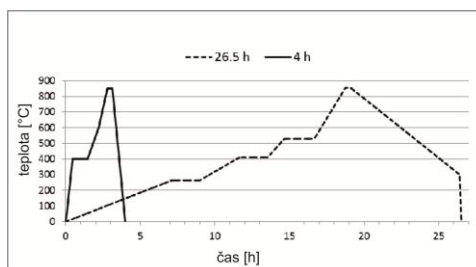
zariadenia DMA analýzy pre skúšku ťahom. Ťahová sila pôsobiaca na vzorku stúpala s rýchlosťou 5 N/min. a maximálna sila bola 18 N. Naše merania boli realizované na zariadení DMA Q800.



Obr. 3 Skúška ťahom v zariadení pre DMA analýzu.

### III. PRÍPRAVA VZORIEK

Pre analýzu mechanických vlastností rôznych typov LTCC substrátov boli použité vyššie opísané testovacie metódy. Rozmery vzoriek boli rovnaké pre všetky typy LTCC substrátov pred vypalovacím procesom. Pre naše experimenty boli použité jedno, dvoj, troj, štvor a päťvrstvové vzorky s LTCC, ktoré boli laminované isostatickou metódou. Rozmery vzoriek boli približne 5 x 30 mm. V procese výroby vzoriek boli použité rôzne vypalovacie profily, ako je to uvedené v Tab. 1. Použité vypalovacie profily sú znázornené na Obr. 4. Pre účely DMA analýzy boli vytvorené trojvrstvové vzorky.



Obr. 4 Použité vypalovacie profily pre výrobu vzoriek z LTCC.

Tabuľka 1  
 Podmienky výpalu LTCC substrátov

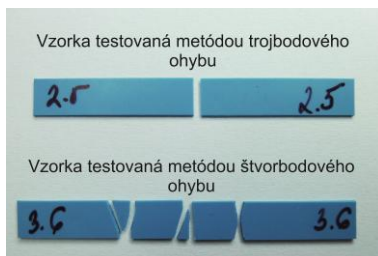
Typ LTCC substrátu	Trvanie vypalovacieho procesu		
	4 h	2 x 4 h	26.5 h
DuPont 951	X	X	X
DuPont 9K7			X
MuRata LFC			X

### IV. ANALÝZA NAMERANÝCH DÁT

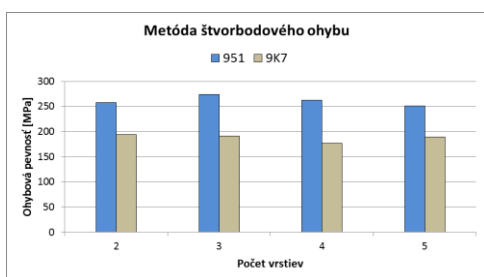
Táto práca opisuje a porovnáva metódy trojbodového a štvorbodového ohybu z hľadiska ich vhodnosti na analýzu mechanických vlastností krehkých materiálov ako sú LTCC substráty. Obr. 6 ukazuje, že metóda štvorbodového ohybu ponúka meranie ohybovej pevnosti materiálu s menšími odchýlkami, pričom trojbodový ohyb poskytuje výsledky s vyššou presnosťou v porovnaní s hodnotami udávanými výrobcami jednotlivých LTCC substrátov. Obr. 6 ďalej



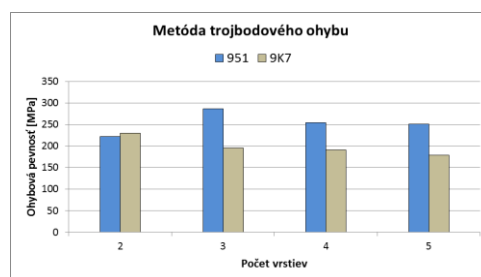
ukazuje, že najvyššie hodnoty ohybovej pevnosti boli namerané pri trojvrstvových vzorkách LTCC DuPont 951. Naše experimenty ukázali, že metóda štvorbodového ohybu ponúka meranie mechanických vlastností materiálu s lepším rozložením mechanického pnutia v objeme materiálu. Táto metóda je vhodná na meranie húževnatosti keramických materiálov. Trojbodový ohyb je vhodný na analýzu tvrdosti materiálu. Táto testovacia metóda sústreďuje mechanické pnutie do jedného bodu v materiáli a následne iniciuje lom v tomto bode. Obr. 5 znázorňuje rozdiel v distribúcii mechanického pnutia v materiáli v prípade trojbodového a štvorbodového ohybu.



Obr. 5 Rozloženie mechanického pnutia v materiáli vplyvom trojbodového a štvorbodového ohybu.

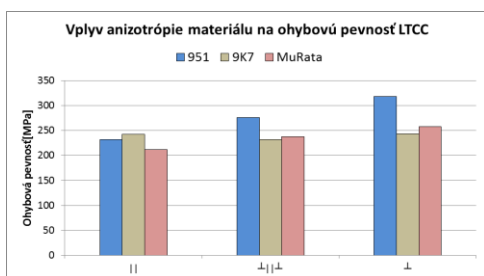


Obr. 6 Vplyv počtu vrstiev na ohybovú pevnosť LTCC meraná metódou štvorbodového ohybu.

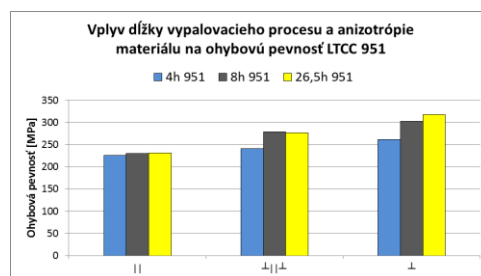


Obr. 7 Vplyv počtu vrstiev na ohybovú pevnosť LTCC meraná metódou trojbodového ohybu.

Anizotropia materiálu má výrazný vplyv na ohybovú pevnosť LTCC substrátu. Experimenty ukázali, že pri všetkých typoch LTCC substrátov bola ohybová pevnosť najvyššia, keď pásiky LTCC boli vyrezané kolmo na smer liatia surovej keramiky LTCC. Naše experimenty boli realizované na rôznych kombináciách smerov keramiky: tri paralelné pásiky ||, kombinácia paralelných a kolmých pásikov ⊥ || ⊥ a tri kolmé pásiky ⊥ (Obr. 8).



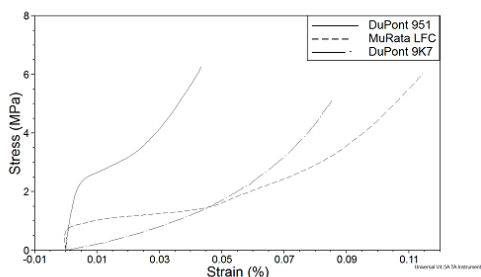
Obr. 8 Vplyv anizotropie materiálu na ohybovú pevnosť troch typov LTCC.



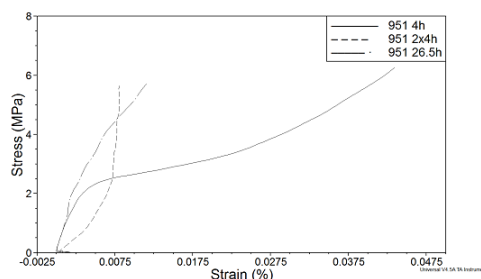
Obr. 9 Vplyv dĺžky výpalu a anizotropie materiálu na ohybovú pevnosť LTCC 951.

Nameraná ohybová pevnosť kolmých trojvrstvových vzoriek metódou trojbodového ohybu bola 320 MPa pre DuPont 951, 243 MPa pre DuPont 9K7 a 257 MPa pre MuRata LFC. Tieto hodnoty sa kompletne zhodujú s hodnotami uvádzanými výrobcami jednotlivých substrátov. Obr. 9 znázorňuje vplyv dĺžky výpalu a anizotropie materiálu na ohybovú pevnosť LTCC

substrátu DuPont 951. Výsledky ukazujú, že ohybová pevnosť LTCC 951 stúpla po 26.5 h vypalovacom procese o 40 %. Obr. 10 a Obr. 11 predstavujú výsledky DMA analýzy. Z Obr. 10 je zrejmé, že najväčšiu pevnosť má substrát DuPont 951. Na druhej strane, DuPont 9K7 a MuRata LFC majú vyššiu hodnotu húževnatosti. DMA analýza ukázala, že opakovaný výpal zvýšil ohybovú pevnosť LTCC, ako to bolo ukázané aj v prípade ohybových testov.



Obr. 10 Porovnanie troch typov LTCC substrátov pomocou DMA analýzy.



Obr. 11 Vplyv vypalovacieho procesu substrátu DuPont 951 na mechanické vlastnosti.

## V. ZÁVER

Výsledky analýzy mechanických vlastností LTCC substrátov ukázali, že metóda štvorbodového ohybu indikuje lepšie rozloženie pnútia v materiáli. Z toho dôvodu, namerané hodnoty ohybovej pevnosti materiálu touto metódou boli nižšie ako hodnoty získané metódou trojbodového ohybu.

Metóda štvorbodového ohybu je vhodná na analýzu húževnatosti materiálu. Trojbodový ohyb inicializuje lom v materiáli v jednom bode.

Odlíšne podmienky výpalu majú značný vplyv na mechanické vlastnosti LTCC substrátov. Naše experimenty ukázali, že opakovaný výpal zvýšil ohybovú pevnosť substrátu DuPont 951 o 10 % a netypický vypalovací proces v trvaní 26.5 h zvýšil ohybovú pevnosť o 40 %. Ohybová pevnosť LTCC substrátov bola vyššia, keď pásiky LTCC boli vyrezané kolmo na smer liatia surovej LTCC. Takto namerané hodnoty ohybovej pevnosti LTCC substrátov boli v korelácii s hodnotami udávanými výrobcami LTCC.

DMA analýza ukázala, že DuPont 951 má najvyššiu mechanickú pevnosť a DuPont 9K7 a MuRata LFC majú vyššiu húževnatosť. Opakovaný výpal znížil húževnatosť substrátu DuPont 951. DMA analýza je vhodná metóda na meranie húževnatosti materiálov na keramickej báze v pružnej oblasti. Táto práca dokazuje, že výsledky DMA analýzy sú v korelácii s výsledkami ohybových skúšok.

## POĎAKOVANIE

Táto práca vznikla na základe podpory projektu VEGA: 1/0218/13.

## REFERENCIE

- [1] W. D. Callister, "Fundamentals of Material Science and Engineering: An Integrated Approach", 3<sup>rd</sup> edition, Wiley 2008, ISBN: 978-0-470-12537-3.
- [2] M. F. H. Wolff, V. Salikov, S. Antonyuk, S. Heinrich, G. A. Schneider, "Three-dimensional discrete element modeling of micromechanical bending tests of ceramic-polymer composite materials", *Powder Technology* 248, 2013, pp. 77-83.
- [3] F. Mujika, "On the difference between flexural moduli obtained by three-point and four-point bending tests", *Polymer Testing* 25, 2006, pp. 214-220.
- [4] I. Petruscu, C. Mohora, C. Ispas, "The determination of young modulus for CFRP using three point bending tests at different span lengths", *U. P. B. Sci. Bull., Series D*, Vol. 75, Iss. 1, 2013, pp. 121-128.
- [5] K. P. Menard, "Dynamic Mechanical Analysis, A practical introduction", 1<sup>st</sup> edition, USA, CRC Press 1999, ISBN: 0-8493-8688-8.
- [6] A. Pietriková, S. Slosarčík, "Mechanical properties of bent modules based on LTCC", *Iranian Journal of Electrical and Computer Engineering*, Vol. 1, No. 1, 2002, pp. 47-50.

# Design of Low Pass Filter Based on Substrate Murata LFC for Ultra Wide-Band Sensor System

*Kornel RUMAN, Alena PIETRIKOVÁ, Igor VEHEC, Tibor Rovenský*

Department of Technologies in Electronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

kornel.ruman@tuke.sk, alena.pietrikova@tuke.sk, i.vehec@tuke.sk,  
tibor.rovensky@tuke.sk

**Abstract** —The aim of this paper lies in a design, simulation and experimental testing of microstrip stepped-impedance Low-Pass (LP) filters with a cut-off frequency of 2.5 GHz. This paper demonstrates the design of LP filter realised by the HyperLynx 3D EM Designer from Mentor Graphics®. The paper presents simulated and measured results of mentioned filters' scattering parameters (S parameters) based on Low Temperature Co-fired Ceramic (LTCC) substrate Murata LFC. This paper assesses the suitability of LTCC substrate Murata LFC for the production of the microstrip stepped-impedance LP filter for high frequency (HF) area. Accent of this article was to demonstrate the compatibility of LTCC substrate Murata LFC with conductor system LL612 Ag designed for GreenTape 9K7. The developed microstrip stepped-impedance LP filters should be used as a part of I – Q (In-phase Quadrature) demodulator presented in [1].

**Keywords** — Murata LFC, Low-Pass, LTCC, HF

## I. INTRODUCTION

The rapidly growing field of UWB applications in various areas pushes the requests for new enhanced UWB radar systems. One of the very promising solutions for the UWB device realization is based on so called M-sequence approach, where the operation is based on a special type of the M-Sequence. This UWB sensor system is simply expandable with I – Q demodulator on side of receiver and mixer on the transmitting side [1]. I – Q demodulation is very useful if we need to obtain both the magnitude and phase of the received signal. Mentioned UWB sensor system operates in frequency band from DC to 2 GHz, for that reason it is necessary to filter signals at I – Q demodulator outputs by the LP filters. This will eliminate undesirable products after mixing, or other spurious signals from HF bands, which would negatively affect the digitization and further processing of the useful signal.

Nowadays, filtering in I – Q demodulator is provided by commercial fabricated LP LTCC filter (LFCN-3400+) from Mini-Circuits Company [1]. However, this filter does not achieve the required properties (cut off frequency 2.5 GHz and attenuation -40 dB in stop band). Commercially fabricated LP filter have pass band from DC to 4.1 GHz [3] which is 2.1 GHz more than is required. Attenuation of commercial fabricated LP filter (LFCN-3400+) in stop band is only -20 dB which is -20 dB less than is required. This unsuitable properties of commercially fabricated LP filter causes aliasing and leads to unobserved Nyquist sampling theorem (if a function  $x_t$  contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced  $1/2B$  seconds apart).

Advances in CAD (Computer-Aided Design), such as full-wave EM (Electromagnetic) simulators did coup in design of filters. In this paper is use for design of the microstrip filter software HyperLynx 3D EM Designer, which facilitated the design of microstrip LP filter based on LTCC.

## II. FILTER DESIGN

### A. Substrates

In the HF area, the dielectric properties of substrate have a major impact on the quality and stability of the filter. Each substrate at the market has various parameters. This is the reason

why not every substrate is proper for UWB area and choosing the appropriate substrates should pay close attention. Those with lower dielectric constants are more suitable for lower frequency applications. The FR4 (Flam Resistant based on epoxy and glass – FR4) laminate is the most common electronic carrier for circuits operating below the microwave frequency. Beyond 1 GHz, the laminate properties (like  $\epsilon_r$ ,  $\tan \delta$ , roughness, etc...) of the FR4 become nonlinear. It becomes difficult to precisely characterize the FR4 at the desired microwave frequency. Beyond 3 GHz, the use of the FR4 is not recommended anymore, in the case of unacceptable attenuation. Significant parameter in HF area is the value of dissipation factor  $\tan(\delta)$  [4].

For correct design and simulation is necessary to know the parameters of the used substrate. As substrate we chose LTCC substrate Murata LFC. LTCC from Murata offer solutions to many of the problems facing designers today. One of the LTCC material systems Murata offers is the LFC® (Low-temperature Fireable Ceramics). This material system has been widely accepted for automotive applications and has established world-wide reputations for its superior characteristics. Our advanced sintering processes create significantly higher accuracy with all the benefits that implies. The Murata LFC comprises a complete cofireable family of Au and Ag metallization, buried passives, and encapsulants. Murata LFC is available in multiple thicknesses and is designed for use as an insulating layer in multichip modules, single chip packages, ceramic printed wiring boards and RF modules [5]. Parameters of used substrate are in Table 1.

Table 1.  
 Specification of used LTCC substrate Murata LFC

Properties	Units	Typical Value
Layer Thickness after firing	$\mu\text{m}$	50, 100, 130 and 160
X, Y Shrinkage	%	16
Z Shrinkage	%	13
Density	$\text{g/cm}^3$	2.9
TCE (25 to 300 °C)	$\text{ppm}/^\circ\text{C}$	5.5
Flexural strength	(MPa)	270
Thermal conductivity	(W/mK)	2
Dielectric constant (6 GHz)	-	7.55
Loss Tangent (6 GHz)	-	0.0045

Low dielectric loss as typical property of LTCC substrate Murata LFC allows using them in many applications where at high operating frequency are conventional laminates circuit boards and LTCC ceramics limited.

### B. Metallization

For design and simulation of microstrip filter we used two different conductor systems: Based on Ag (DuPont LL612) as well as based on AgPd (DuPont LF171). Conductor system LF171 is compatible with LTCC substrate Murata LFC. For considering the compatibility of LTCC substrate Murata LFC with conductor system designed for GreenTape 9K7 we chose conductor system LL612 Ag. Basic specification of both conductor systems is in the Table 2.

Table 2.  
 Specification of used metallization [6,7]

Properties	Value	Value
Conducting pastes	LL612 Ag	LF171 AgPt
Viscosity (Pa.S)	150 - 300	180 - 280
Dried Line Resolution ( $\mu\text{m}$ ) line / space	125 / 125	125 / 125
Fired Thickness ( $\mu\text{m}$ )	9 - 12	13 - 19
Fired Resistivity ( $\text{m}\Omega / \text{sq}$ )	$\leq 3$	$\leq 2.5$
Substrate compatibility	9K7	9K7/951/Murata

Resistivity of used conductor systems has major impact on accuracy of microstrip LP filter. We choose these conductor systems for its low value of resistivity and fine line resolution. Low value of resistivity allows these systems to be used in applications where is requiring excellent conductivity.

C. Simulation

Advances in CAD (Computer-Aided Design), such as full-wave EM (Electromagnetic) simulators did coup in design of filters. For design of the resonators and calculation of the coupling coefficients of microstrip LP filters was used HyperLynx 3D EM Designer software. Based on result of comparison of various possibilities for the shape of microstripe LP filter that offer CAD software we decided for stepped-impedance type of structure. The Fig. 1 shows the structure of five-pole stepped-impedance microstrip LP filter (with its dimensions), which used a cascaded configuration of alternating high and low impedance transmission lines [8].

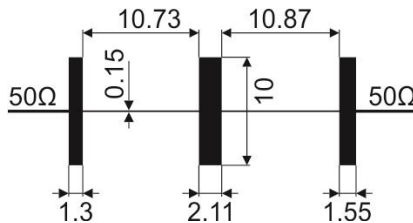


Fig. 1 A layout of the designed microstrip stepped-impedance LP filter for LTCC substrate Murata LF (unit: mm).

The simulation must calculate with shrinkage of LTCC ceramics (Table 1) and conductive paste because Murata doesn't have zero shrinkage after firing process in axis x, y and z. Shrinkage of ceramics plays a major role since results of simulations are calculated for dimension of LP filter after firing.

The Fig. 2 shows simulated result of S parameters after planar EM analysis of microstrip stepped-impedance LP filter. In this simulation, we verified the suitability of the ceramic material Murata and both conductor systems DuPont LL612 and LF171 for the realization of this filter.

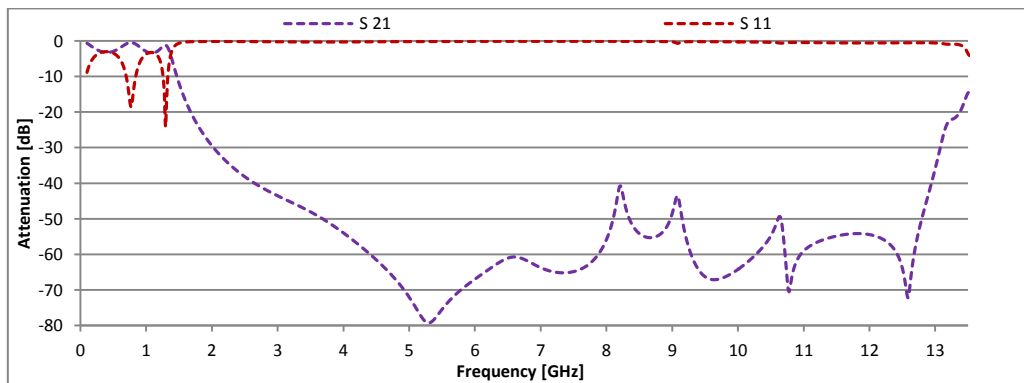


Fig. 2 Full-wave EM simulated performance of the microstrip stepped-impedance LP filter.

Forward transmission coefficient (S21) of designed microstrip stepped-impedance LP filter does not exceed our required minimum value of -40 dB for attenuation in stop band up to 13 GHz (Fig. 2). Designed microstrip stepped-impedance LP filter have pass band from DC to 2.5 GHz what is exactly as we required.

III. RESULTS

A. Filter construction

The final prototype of the microstrip stepped-impedance LP filter including SMA (Sub-Miniature version A) connectors is shown in Fig. 3.

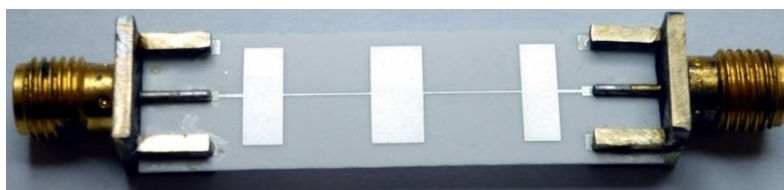


Fig. 3 Final prototype of the microstrip stepped-impedance LP filter.

The microstrip stepped-impedance LP filters were produced by conventional screen printing technology. Quality and thickness of the print depends on many screen printing factors (printing pressure, printing speed, thickness of the emulsion, characteristic of the cloth, etc.). Very important role plays firing conditions as well. The roughness on boundary of transmission lines influences quality of transmitted signals in HF area. That's because the roughness can cause reflection of transmitted signals what is leading to the signal attenuation. To achieve fine line of BP structure characterized by homogeneity of thickness, without roughness of the surface and line edges we apply 325 mesh cloth screens in combination with capillary film thickness of 29  $\mu\text{m}$  were used.

### B. Measurement of $S$ parameters

The forward transmission coefficient and return losses were measured using the Rohde & Schwarz vector network analyzer. Comparisons of simulated and measured forward transmission coefficient of three microstrip stepped-impedance LP filters are showed in the Fig. 4.

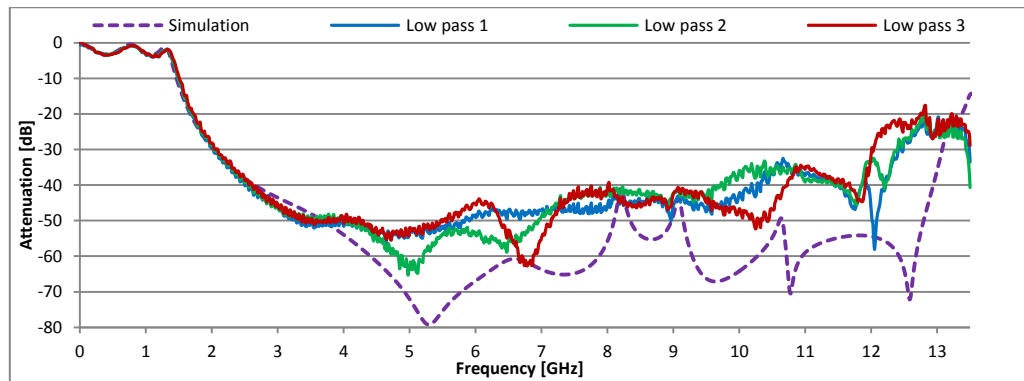


Fig. 4 Comparison of simulated and measured forward transmission coefficient of microstrip stepped-impedance LP filters.

The forward transmission coefficient of the microstrip stepped-impedance LP filters made from LTCC substrate Murata LFC doesn't match with initial conditions (attenuation -40 dB in stop band) in the whole frequency range (from DC to 13.5 GHz). We can observe peak on the measured LP filters' performances at 10.5 GHz. Even peaks exceed our required minimum value of -40 dB for attenuation in stop band the transmission characteristics of measured LP filters are still acceptable. On the other hand the cut off frequency of measured microstrip stepped-impedance LP filters is exactly at required value 2.5 GHz. It is possible to use this filter as antialiasing LP filter in UWB devices such as IQ demodulator presented in [1] or part of system for Through Wall Moving Target Tracking by M-sequence UWB Radar presented in [9].

Comparisons of simulated and measured return losses of three microstrip stepped-impedance LP filters are in the Fig. 5.

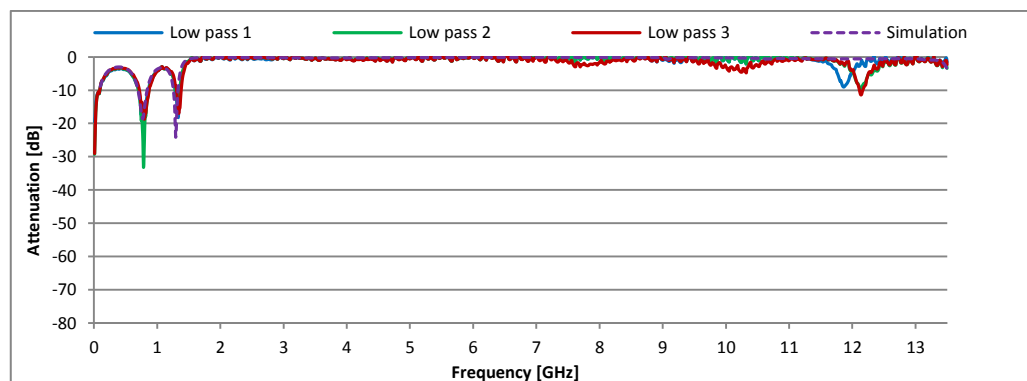


Fig. 5 Comparison of simulated and measured return losses of microstrip stepped-impedance LP filters.

The Fig. 5 shows that transmission characteristics of microstrip stepped-impedance LP filters match the simulation in the whole frequency range. It is not possible to observe significant

divergences between measured and simulated characteristics of microstrip stepped-impedance LP filters. Transmission characteristics of measured LP filters meet with our requirements from DC to 13.5 GHz.

#### IV. CONCLUSION

The microstrip stepped-impedance LP filters with cut off frequency 2.5 GHz and a minimum attenuation -40 dB in stop band were designed, simulated and experimental tested. The layouts of LP filters were created by the HyperLynx 3D EM Designer (from Mentor Graphics) and simulated with full-wave electromagnetic simulation. Comparisons between simulated and measured S parameters of designed microstrip stepped-impedance LP filters were presented. We point out that conductor system designed for GreenTape 9K7 (DuPont LL612) is compatible with LTCC substrate Murata LFC. We verified the suitability of the LTCC material Murata LFC for the production of the microstrip stepped-impedance LP filters for high frequency area. Transmission characteristics of microstrip LP filters are acceptable for use in the UWB sensor system as an antialiasing LP filters for I – Q demodulator.

#### ACKNOWLEDGMENT



Paper is the result of the Project implementation: University Science Park TECHNICAL COM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia/This project is being co-financed by the European Union.

This paper was developed with support of the project: Effects of behaviour of multilayer modules based on LTCC in the high frequency environment, VEGA: 1/0218/13.

#### REFERENCES

- [1] Žiga M., Liptaj M., "Design of the I-Q Demodulator for UWB Applications", SCYR 2012 - 12th Scientific Conference of Young Researchers - FEI TU of Košice, 2012, Košice
- [2] ECC Decision (06)04, "The harmonised conditions for devices using Ultra-Wideband (UWB) technology in bands below 10.6 GHz," 2006. [Online]. Available: <http://www.erodocdb.dk/docs/doc98/official/pdf/ECCDEC0406.pdf>.
- [3] Ceramic Low Pass Filter, "Ceramic Low Pass Filter LFCN-3400 +, 50Ω, from DC to 3400 MHz," 2014. [Online]. Available: <http://217.34.103.131/pdfs/LFCN-3400+.pdf>.
- [4] Nikholas G. Toledo, "Practical Techniques for Designing Microstrip Tapped Hairpin Resonator Filters on FR4 Laminates", in Communications Engineering Division, Advanced Science and Technology Institute, Diliman Quezon City, Philippines 1101
- [5] Technology Focus, "Using LFC Technology, Murata Makes Automotive Ceramic Substrates," 2014. [Online]. Available: <http://www.murata.com/articles/ta04e2.pdf>
- [6] Dupont Microcircuit Materials, "Technical Data Sheet Dupont LL612," 2014. [Online]. Available: [http://www2.dupont.com/MCM/en\\_US/assets/downloads/prodinfo/LL612.pdf](http://www2.dupont.com/MCM/en_US/assets/downloads/prodinfo/LL612.pdf)
- [7] Dupont Microcircuit Materials, "Technical Data Sheet Dupont LF171," 2014. [Online]. Available: [http://www2.dupont.com/MCM/en\\_US/assets/downloads/prodinfo/LF171.pdf](http://www2.dupont.com/MCM/en_US/assets/downloads/prodinfo/LF171.pdf)
- [8] Jia-Sheng Hong, "Microstrip Filters for RF/Microwave Applications," John Wiley & Sons, Inc., Hoboken, 2nd edition, New Jersey, 2011, 655.
- [9] D. Urdzik, D. Kocur, and J. Rovnáková, "Detection of multiple targets with enhancement of weak UWB radar signals for the purposes of through wall surveillance," in Applied Machine Intelligence and Informatics (SAMI), 2012 IEEE 10th International Symposium on, jan. 2012, pp. 137 –142.

# Simulation of a Two-Phase Induction Motor

<sup>1</sup>Želmíra FERKOVÁ, <sup>2</sup>Ján KAŇUCH

<sup>1,2</sup>Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>zelmira.ferkova@tuke.sk, <sup>2</sup>jan.kanuch@tuke.sk

**Abstract** — This paper addresses modeling a two-phase induction motor. Based on geometric dimensions, the model was created in the ANSYS/Maxwell program. Within the same program, the transient process when the motor starts up was also calculated. Another model was made in the MATLAB/Simulink program on the basis of a mathematical description of the induction motor. The waveforms of the currents and torques of the motor are compared. The model created in MATLAB/Simulink is compared with measurements. A symmetrical two-phase power supply was used for the measurements.

**Keywords** — two-phase induction motor, MATLAB/Simulink, dynamic model

## I. INTRODUCTION

Three-phase asynchronous machines are powered by three-phase symmetrical voltage. The windings of the motors are arranged within the stator and are mutually shifted by 120 electrical degrees. These machines operate with circular magnetic fields.

Single-phase motors have two windings (phases) in their stators, which are mutually shifted by 90 electrical degrees. One of the windings is primary, and the other is auxiliary. Both phases are powered by a single-phase network. The time shift in the auxiliary winding is usually due to a capacitor connected in series with the auxiliary winding. The magnetic field formed in this way is elliptical. The torque of the motor consists of a forward torque and a backward torque. If both the main windings and the auxiliary windings of a single-phase motor are identical and are powered by two-phase voltage, a circular magnetic field will be developed within the machine.

This paper compares two-phase motor models, one developed within MATLAB/Simulink and the other in the ANSYS/Maxwell program. The model in the ANSYS/Maxwell program is based on the simplified geometric shape of a real motor, and takes into account the non-linearity of the magnetic materials that are used. The MATLAB/Simulink model works with constant parameters that were obtained by measurements performed on a real motor with the same main dimensions.

## II. MATHEMATICAL MODEL

A model that takes into account asymmetrical windings was used to simulate a two-phase motor in the MATLAB/Simulink program [1]. The varying number of turns in the phases is reflected by parameter “a”. The system of fixed stator coordinates  $\alpha$ - $\beta$  is applied. In this specific case, the asymmetry of the parameters of the motor is due to the winding production. There is an identical number of turns in the main phase and in the auxiliary phase.

The components of the stator and the rotor voltage of a two-phase induction motor can be expressed as follows:

$$u_{s\alpha} = R_{s\alpha} i_{s\alpha} + \frac{d\psi_{s\alpha}}{dt}, \quad (1)$$

$$u_{s\beta} = R_{s\beta} i_{s\beta} + \frac{d\psi_{s\beta}}{dt}, \quad (2)$$

$$u_{r\alpha} = R_{r\alpha} i_{r\alpha} + \frac{d\psi_{r\alpha}}{dt} + a\omega_{r\alpha} \psi_{r\beta}, \quad (3)$$



$$u_{r\beta} = R_{r\beta}i_{r\beta} + \frac{d\psi_{r\beta}}{dt} + a\omega_r\psi_{r\alpha}, \quad (4)$$

where  $a$  is the ratio of the main and auxiliary winding turns.

The components of the stator and rotor flux linkages are:

$$\begin{aligned} \psi_{s\alpha} &= L_{s\alpha}i_{s\alpha} + L_{m\alpha}i_{r\alpha}, & \psi_{s\beta} &= L_{s\beta}i_{s\beta} + L_{m\beta}i_{r\beta} \\ \psi_{r\alpha} &= L_{m\alpha}i_{s\alpha} + L_{r\alpha}i_{r\alpha}, & \psi_{r\beta} &= L_{m\beta}i_{s\beta} + L_{r\beta}i_{r\beta}. \end{aligned} \quad (5)$$

The components of the stator and rotor currents are:

$$\begin{aligned} i_{s\alpha} &= \frac{L_{r\alpha}\psi_{s\alpha} - L_{m\alpha}\psi_{r\alpha}}{L_{s\alpha}L_{r\alpha} - L_{m\alpha}^2}, & i_{s\beta} &= \frac{L_{r\beta}\psi_{s\beta} - L_{m\beta}\psi_{r\beta}}{L_{s\beta}L_{r\beta} - L_{m\beta}^2} \\ i_{r\alpha} &= \frac{L_{s\alpha}\psi_{r\alpha} - L_{m\alpha}\psi_{s\alpha}}{L_{s\alpha}L_{r\alpha} - L_{m\alpha}^2}, & i_{r\beta} &= \frac{L_{s\beta}\psi_{r\beta} - L_{m\beta}\psi_{s\beta}}{L_{s\beta}L_{r\beta} - L_{m\beta}^2}. \end{aligned} \quad (6)$$

The electromagnetic torque produced by the machine is given by equation (7), and the dynamic equation (8) was used:

$$m_i = p(L_{m\beta}i_{s\beta}i_{r\alpha} - L_{m\alpha}i_{s\alpha}i_{r\beta}), \quad (7)$$

$$m_i - m_z = \frac{J}{p} \cdot \frac{d\omega}{dt}, \quad (8)$$

where

- $\psi_{s\alpha}, \psi_{s\beta}, \psi_{r\alpha}, \psi_{r\beta}$  - are the stator and rotor flux linkages,
- $i_{s\alpha}, i_{s\beta}, i_{r\alpha}, i_{r\beta}$  - are the stator and rotor currents,
- $L_s, L_r, L_m$  - are the stators, rotor, and magnetizing inductances,
- $R_s, R_r$  - are the stator and rotor resistances,
- $\omega$  - is the electrical rotor angular speed,
- $J$  - is the moment of rotor inertia,
- $m_i$  - is the electromagnetic torque of the motor.

The equivalent circuits [1] representing the two-phase induction motor in the stationary ( $\alpha\beta$ ) reference frame are shown in figure 1.

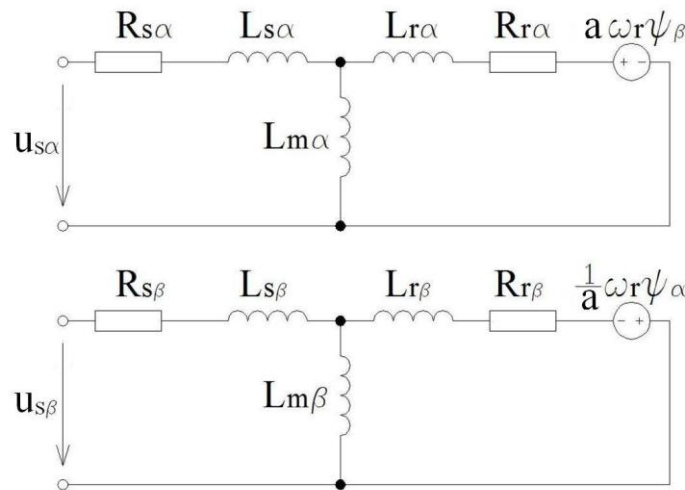


Fig. 1 Equivalent circuit of an asymmetrical two-phase induction motor in the stationary ( $\alpha\beta$ ) reference frame.

### III. MODEL OF THE MOTOR IN ANSYS/Maxwell

Exactly the same motor was simulated in the ANSYS/Maxwell program. We calculated the start-up of the motor when directly connected to the voltage. The magnetic field of the motor was calculated by the finite elements method in 2D (the transient solution). Nonlinear properties of the magnetic circuitry material were taken into account when making the calculations. A model of the motor is shown in Fig. 2.

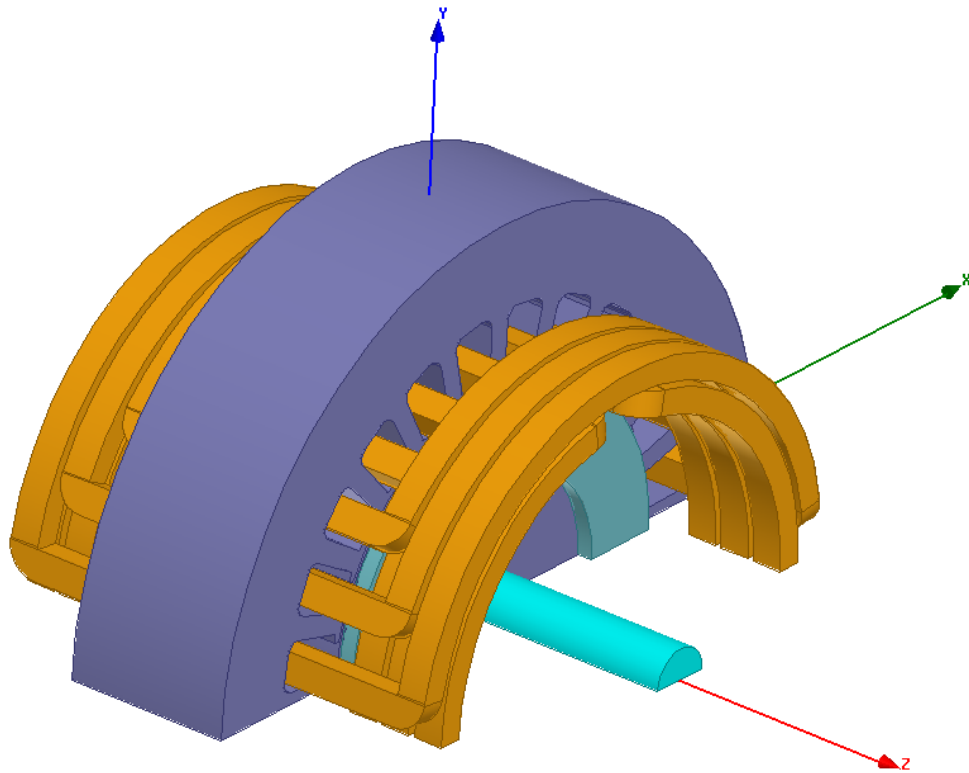
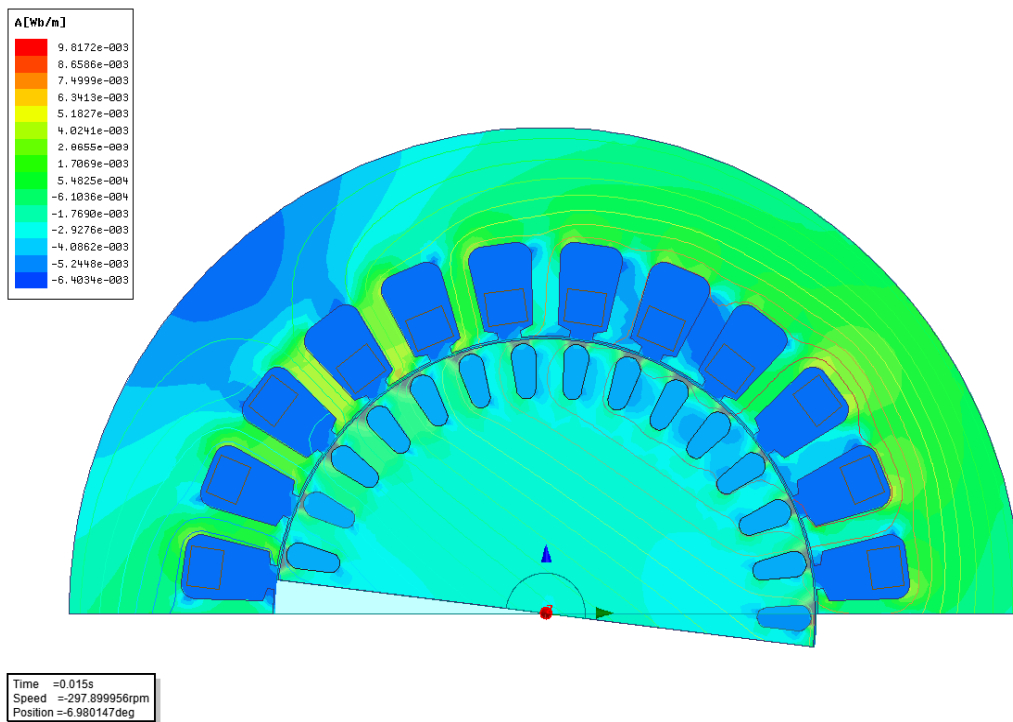


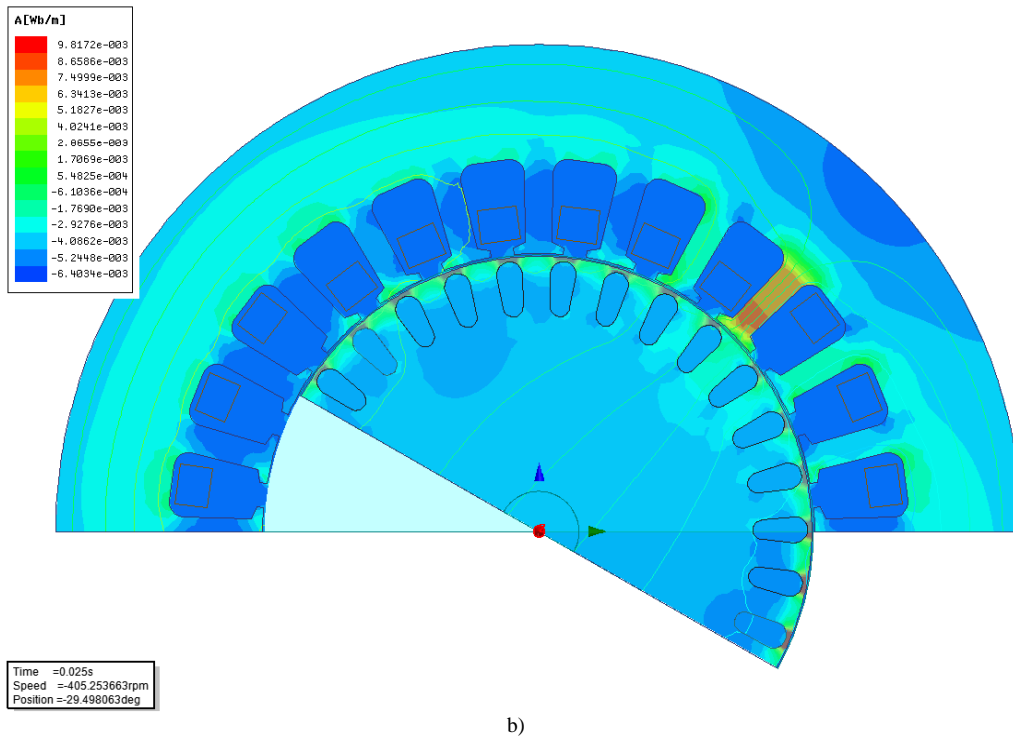
Fig. 2 Model of the two-phase motor.

Fig. 3 shows the waveforms of the magnetic field calculated for time  $t=0,015$  s (Fig. 3a) and for time  $t=0,025$  s (Fig. 3b).

An external excitation circuit was chosen (Fig. 4), where LA and LB are the end leakage inductances. The end leakage inductances were calculated in the ANSYS/Rm-expert program for the type of winding that was used.



a)



b)  
 Fig. 3 Magnetic field of the two-phase motor.

The motor is supplied with two-phase symmetrical voltage (Fig. 5). The calculated waveforms of the currents for individual phases of the inner torque and rpm are demonstrated in Fig. 6, 7 and 8.

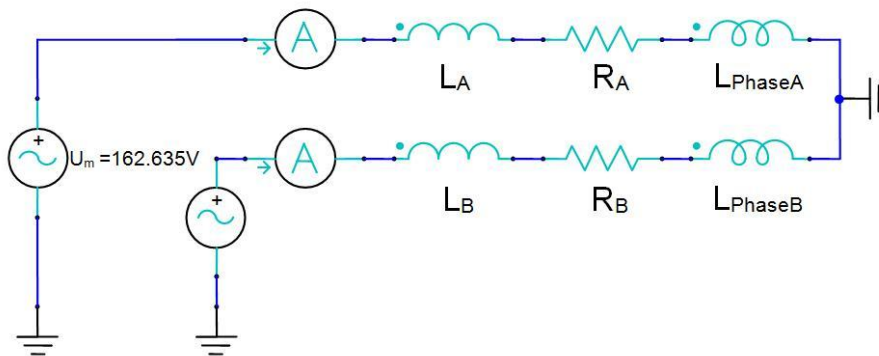


Fig. 4 The external excitation circuit.

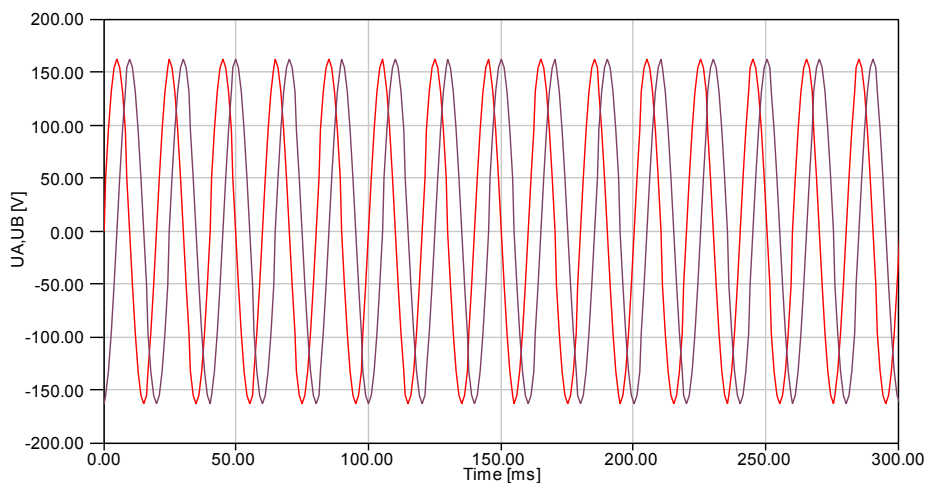


Fig. 5 The two-phase symmetrical voltage.

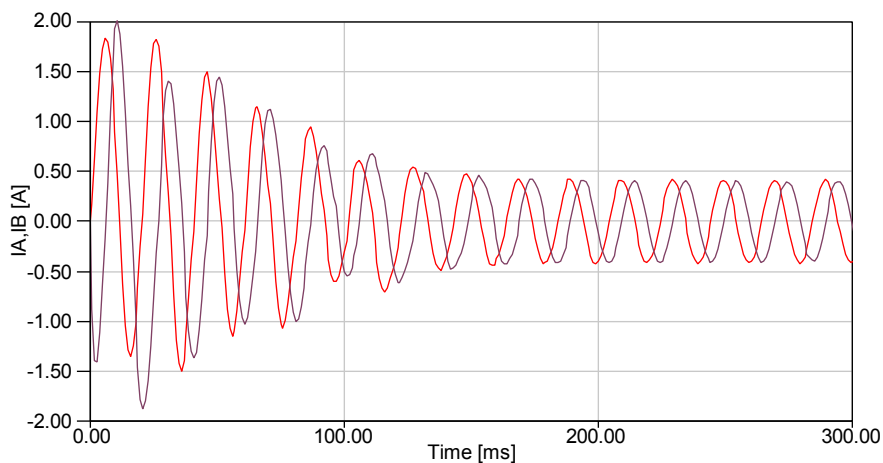


Fig. 6 Calculated waveforms of the motor currents.

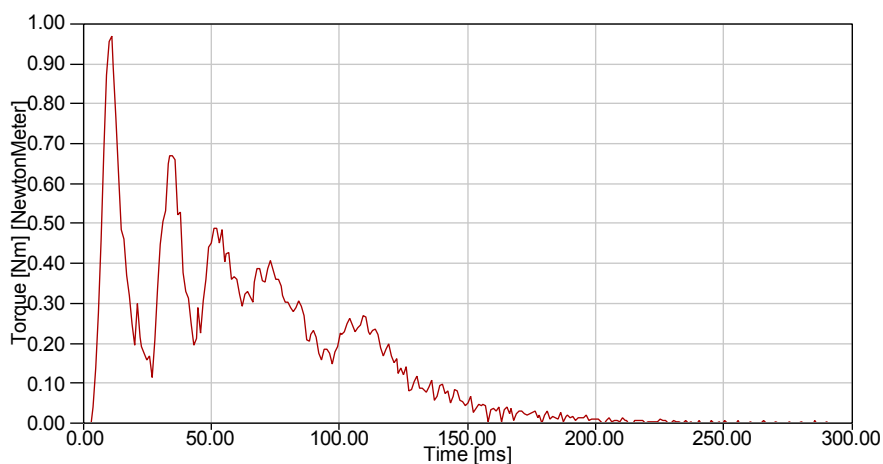


Fig. 7 Calculated waveform of the motor torque.

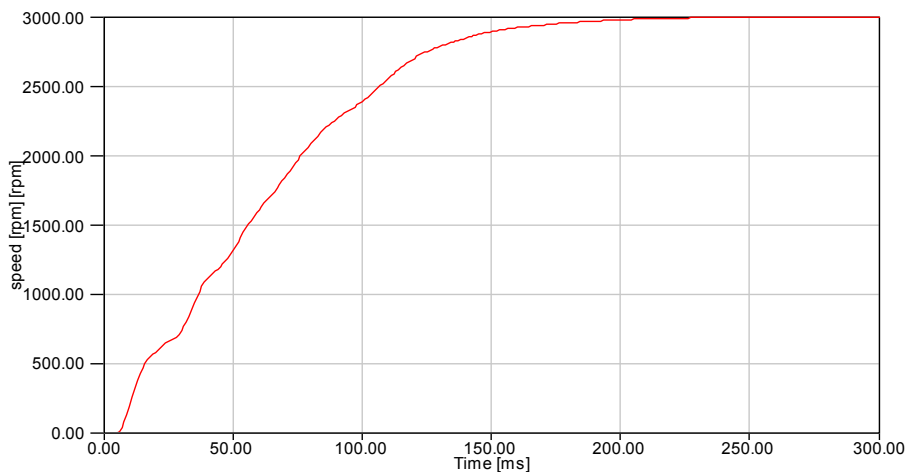


Fig. 8 Calculated waveform of the motor speed.

#### IV. COMPUTER SIMULATION AND EXPERIMENTAL RESULTS

The model of a two-phase asynchronous motor in the MATLAB/Simulink program was created on the basis of equations (1) to (8). The block diagram of the motor is shown in figure 9. The parameters of the motor used for the calculations (resistances, inductances, and the moment of inertia) had been acquired by measurements performed on a specific motor. Other values of the motor parameters were obtained from measurements in the no-load state and in the short-circuit state. The values are shown in table 1.

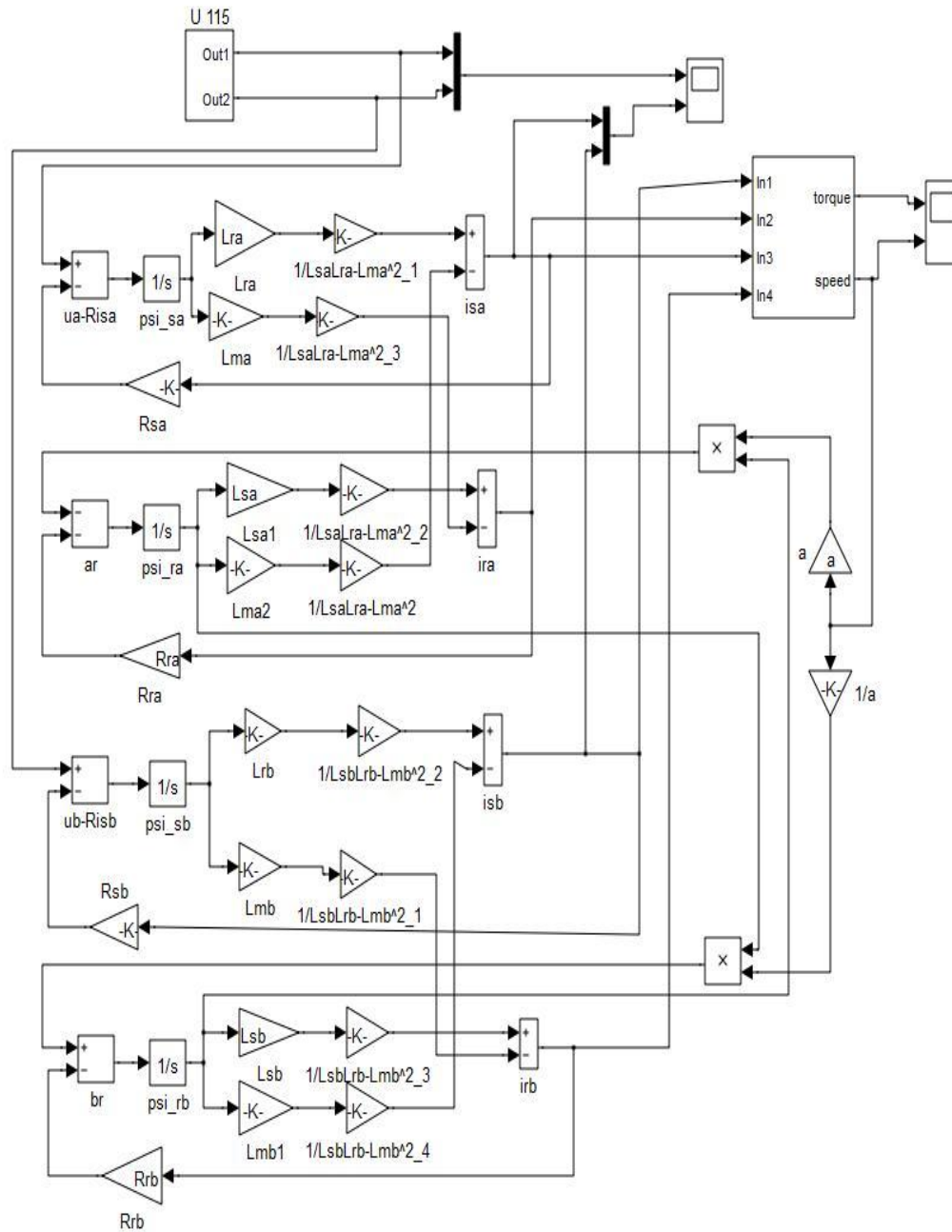


Fig. 9 The scheme of the model for simulation.

Table 1  
 Parameters of the Two-Phase Induction Motor

Parameter	Value	Parameter	Value
$R_{sa}$	30,9 $\Omega$	$R_{ra}$	51 $\Omega$
$R_{sb}$	31,1 $\Omega$	$R_{rb}$	51,35 $\Omega$
$L_{ma}$	1,187 H	$L_{ra}$	1,277 H
$L_{mb}$	1,305 H	$L_{rb}$	1,402 H
$L_{sa}$	1,277 H	J	0,00014 $\text{kg}\cdot\text{m}^2$
$L_{sb}$	1,402 H	f	50 Hz

We simulated the start-up of the motor through direct connection to the supply without loading. The calculated waveform was compared with the results of measurements performed on a real two-phase motor. For the sake of precision, the voltage as read during the measurements was fed into the model (Fig.10). The figure shows that the supply voltage was not perfectly sinusoidal and that the supply unit was relatively soft. The voltage and current values of the motor were recorded using an oscilloscope.

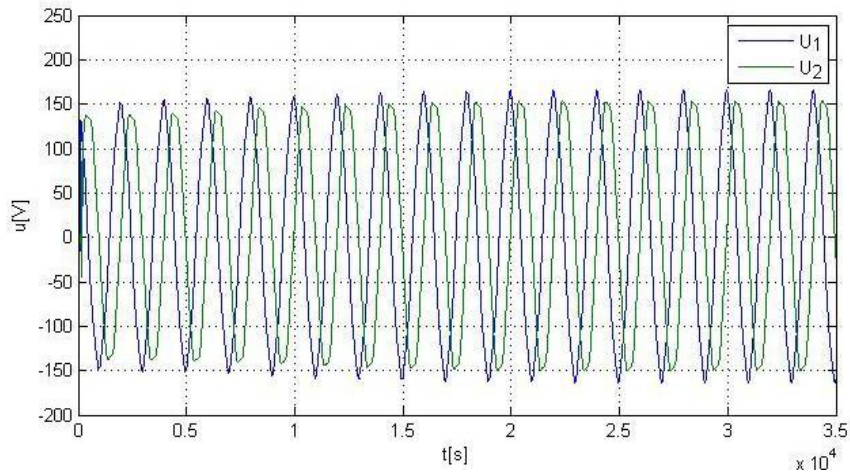


Fig. 10 Simulation of the supply voltage of the two-phase induction motor (first and second phase).

The waveforms of the currents acquired by modeling were comparable with the measurement results. Figure 11 illustrates the calculated waveform of the first phase current and also the measured current waveform. Figure 12 shows both the calculated and the measured waveforms of the second phase current.

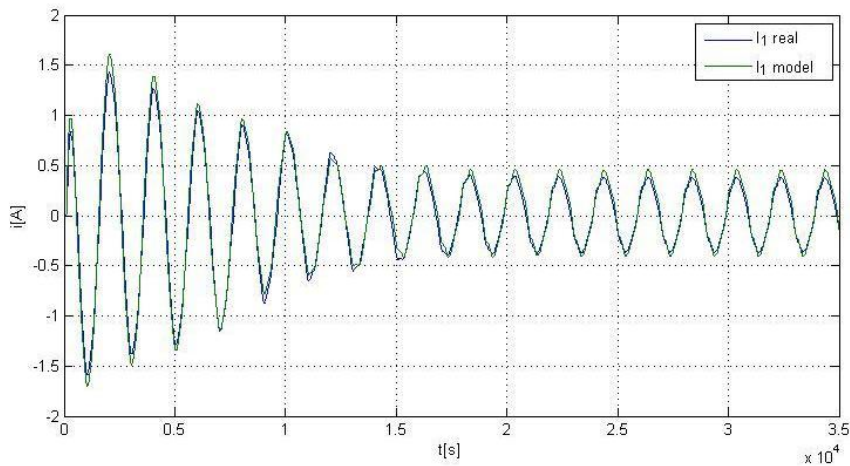


Fig. 11 A comparison of the current in the first phase in the simulation and the measurement of the two-phase induction motor.

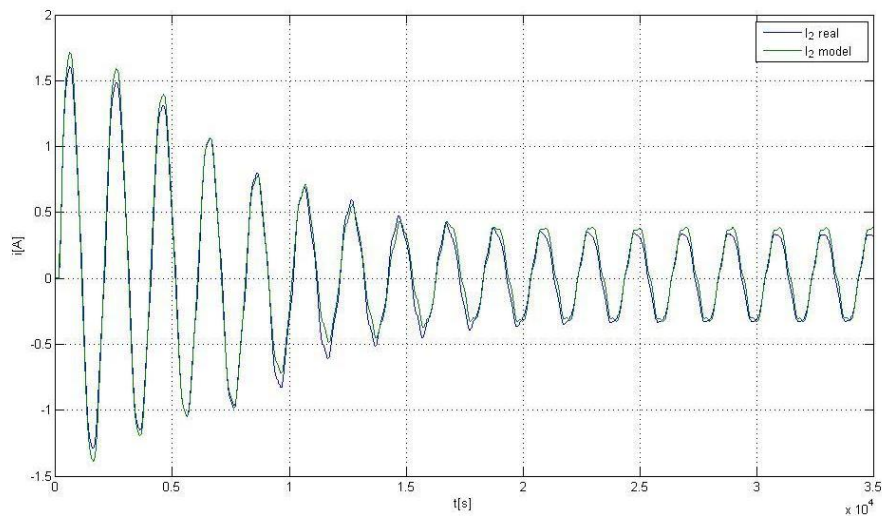


Fig. 12 A comparison of the current in the second phase in the simulation and the measurement of the two-phase induction motor.

Fig. 13 shows the calculated angular speed waveform, and Fig. 14 shows the calculated motor torque waveform.



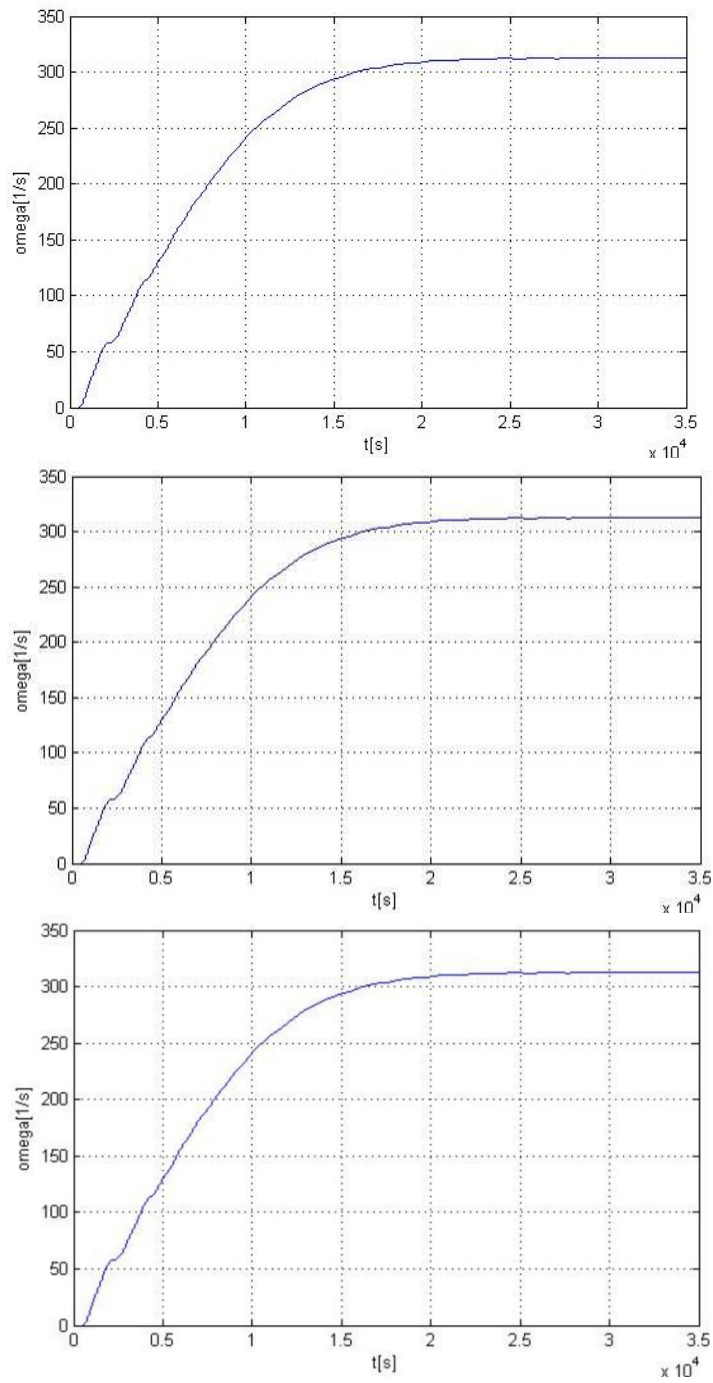


Fig. 13 The rotation speed of the two-phase induction motor.

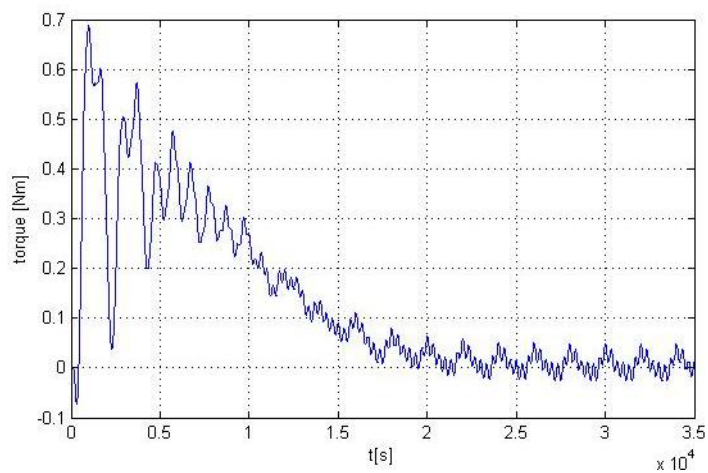


Fig. 14 The calculated electromagnetic torque of the two-phase induction motor for a no-load start.

The revolutions per minute and the torque waveforms of the motor were not measured and are only simulated. The waveforms show a relatively high degree of congruence.

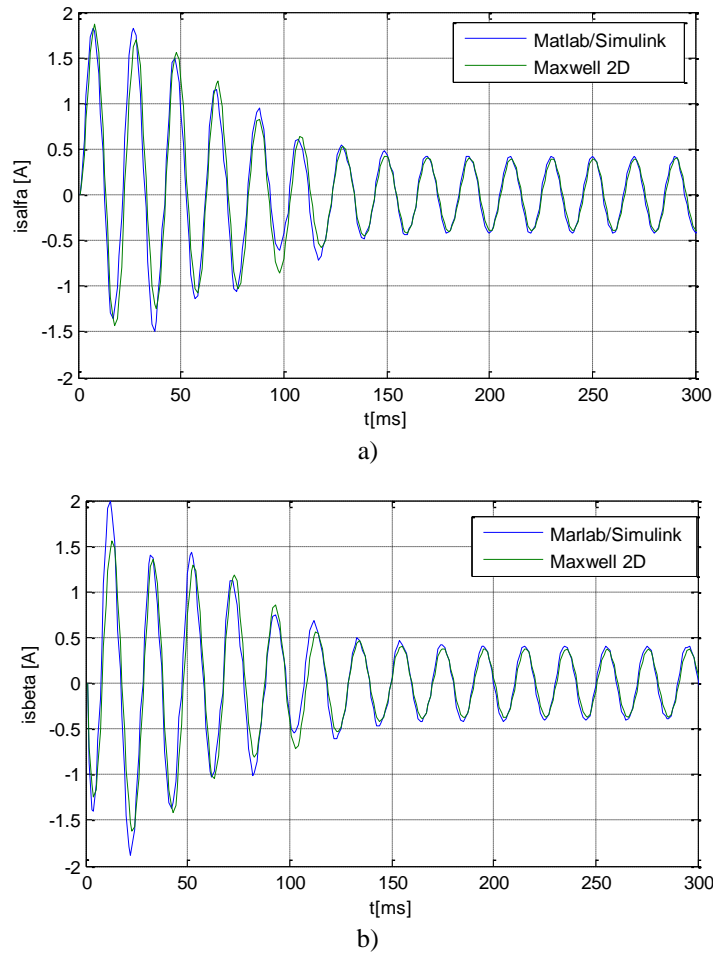


Fig. 15 A comparison of the calculated waveforms of the currents in the MATLAB/Simulink and ANSYS/Maxwell 2D programs.

Figures 15 and 16 compare the calculated waveforms of the currents and the torque (obtained from MATLAB/Simulink and Ansys/Maxwell 2D).

Unlike the waveform in Figs. 10 to 14, in this case the symmetric 2-phase sinusoidal voltage 115 V, 50 Hz with constant amplitude was used. In both cases, start-up by direct connection and without loading was simulated.

Figures 15 a) and b) compare the waveforms of the currents of individual phases ( $I_{S\alpha}$  and  $I_{S\beta}$ ). Good agreement of the current waveforms can be seen in the steady state. Fig. 16 compares the torque waveforms.

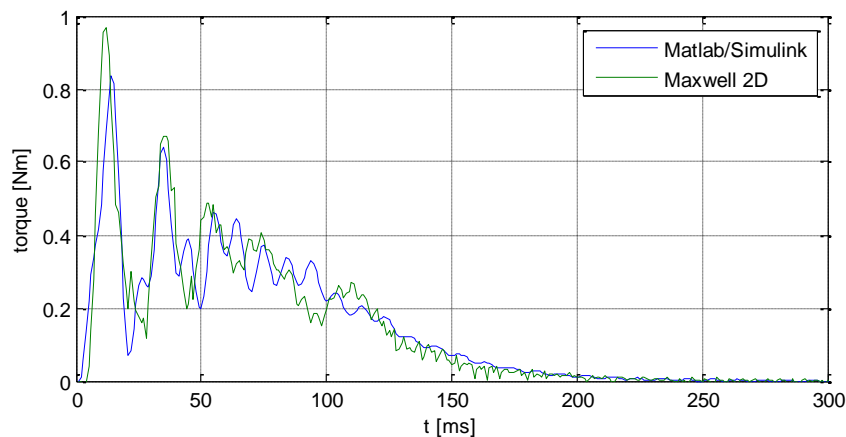


Fig. 16 A comparison of the calculated torque waveforms in the MATLAB/Simulink and ANSYS/Maxwell 2D programs.



The ANSYS/Maxwell program takes into account non-linear properties of the magnetic circuit, and the influence of "the end of the windings" is considered additionally through constant  $L_A$  and  $L_B$  (Fig. 4). The influence of slot skewing is not taken into account. The oscillations in the torque waveforms are due to the imperfect symmetry of the windings. The Maxwell waveform takes into account the non-linearity of the motor parameters, and also the impact of the slots.

## V. CONCLUSIONS

This paper has presented two models of a two-phase motor. A model of the same motor was created in the ANSYS/Maxwell 2D program and in the MATLAB/Simulink program. A comparison was made of the waveforms provided by the two models. The results are not identical. The Simulink model works with constant parameters, whilst ANSYS/Maxwell 2D calculates the magnetic field using the finite element method. ANSYS/Maxwell takes into account the nonlinearity of the magnetic material and also the influence of the slots. Whereas the influence of end of the windings are considered only through the additional inductance and resistance.

In this paper, only the model created in MATLAB/Simulink simulation software has been compared with the measurements. The reason for this was the soft and non-symmetrical power supply (Fig. 7). Since an oscilloscope had recorded the waveform of the supply voltage, the input for the MATLAB/Simulink model used these more precise voltage values. We have compared the waveforms of the current with an identical supply voltage.

When the simulation models (MATLAB/Simulink and ANSYS/Maxwell 2D) were compared, a symmetrical supply voltage with constant amplitude was used.

## ACKNOWLEDGMENT

The financial support of the Slovak Research and Development Agency under the contract No. APVV-0138-10 is acknowledged.

## REFERENCES

- [1] Y. Kumsuwan, W. Srirattanawichaikul, S. Premrudeepreechacharn, *Analysis of two-phase induction motor using dynamic model based on MATLAB/Simulink*. Asian Journal on Energy and Environment 2010, 11(01), pp. 48-59.
- [2] O. Ojo, O. Omozusi, *Parameter Estimation of Single-Phase Induction Machines*. IEEE Industry Applications Conference, vol. 4, September/October 2000, pp. 2280-2287.
- [3] M. E. Popescu, D. Demeter, V. Micu, *Analysis of a voltage regulator for a two phase induction motor drive*. IEEE/IEMDC, May 1999. pp. 658-660.
- [4] G. Craciunas, *Field oriented control of a two phase induction motor*. International Symposium on System Theory, XI Edition, Craiova, 2003, pp. 28-31.
- [5] V. Hrabovcová, P. Rafajdus, M. Franko, *Measuring and modeling of electrical machines*. University of Žilina Press, 2004, Slovakia, (in Slovak).
- [6] N. Abdel-Rahim, A. Shaltout, *Unsymmetrical two-phase induction motor drive with slipfrequency control*. IEEE Trans. Energy Conversion, Vol. 24, No.3, September 2009, pp. 608-616.
- [7] R. Krishnan, *Electric Motor Drives: Modeling, Analysis and Control*. Upper Saddle River, NJ: Prentice-Hall, 2001, p. 350.
- [8] Y. Kumsuwan, S. Watcharin, P. Suttichai, *Analysis of a two-phase induction motor using dynamic model based on MATLAB/Simulink*. Asian Journal on Energy and Environment, 2011, no.4, pp. 2065- 2067.

# Učenie sa chôdze mobilného robota neurónovou sieťou s využitím metódy ZMP

*Ján Vaščák*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

jan.vascak@tuke.sk

**Abstrakt** — Základnou metódou na riadenie pohybu kráčajúcich robotov je prístup založený na hľadani ťažiska bodu z nulovým momentom, tzv. Zero Moment Point (ZMP). Snahou je riadiť robot tak, aby toto bolo premietnuté do plochy jedného z chodidiel robota. Ako to zabezpečiť, je predmetom celej rady prístupov využívajúcich ZMP. V tejto práci sa venujeme využitiu neurónových sietí pri zabezpečení priemetu ZMP do požadovanej polohy. Experimenty boli vykonané na robotovi NAO a sú uvedené a porovnané aj niektoré výsledky.

**Kľúčové slová** — stabilita, humanoidný robot Nao, chôdza

## I. ÚVOD

Rozlišujeme tri kategórie chodiacich dvojnôhých robotov, tzv. humanoidných robotov. Do prvej kategórie patria roboti, ktorých rýchlosť chôdze je príliš malá na to, aby sa robot dostal do stavu, v ktorom by stratil stabilitu pri chôdzi, do stavu nerovnovážnosti. Takýto systém sa dá úplne popísať. Roboti, ktorí majú dynamický pohyb a o niekoľko stupňov voľnosti viac ako roboti v prvej kategórii, patria do druhej kategórie za predpokladu, že majú dostatočné veľké chodidlo. Ich bod nulového ťažiska – Zero moment point (ZMP) je v podporovanom polygóne. V poslednej kategórii sú roboti, ktorí majú len v nohách minimálne 3 - 4 stupne voľnosti, a tým pádom sú aj najzložitejší v otázke riadenia a udržania stability. V prípade, že je v takomto robotovi použitá komplexná metóda udržania stability, je robot schopný dosiahnuť aj pomerne vysokú rýchlosť chôdze [1].

Stabilita humanoidných robotov závisí od veľa atribútov. Môžeme spomenúť napríklad rýchlosť chôdze, počet stupňov voľnosti ako v nohách, tak isto aj v celom tele alebo typ zeme, na ktorej sa robot pohybuje. V prípade vytvorenia algoritmu na stabilitu, ktorý dokáže využiť aj pasívne časti robotov ako napr. senzory alebo kamera, môžeme dosiahnuť ešte lepšie výsledky. Samozrejmosťou je ale úplné poznanie robota, s ktorým pracujeme.

Keď je dvojnôhý robot v pohybe (chôdza), pôsobia naň zotrvačné sily vyvolané gravitáciou a zrýchleniami a spomaleniami zotrvačných síl (inertial force) pri samotnej chôdzi. Výslednica týchto síl sa nazýva celková zotrvačná sila (total inertial force) a má pôsobisko v ťažisku robota. Kontakt robota s podložkou vyvoláva reakčnú silu (floor reaction). V prípade, že robot udržiava perfektnú rovnováhu, sú vektory celkovej zotrvačnej sily a reakcie od podložky v jednej osi. Ak sa tieto sily dostanú mimo svojho osového zarovnania (nerovnosť povrchu), dochádza k strate rovnováhy a generuje sa padacia sila (falling force) [3].

Chôdza dvojnôhých robotov alebo mechanizmov je už dlhé roky pozornosťou výskumníkov a profesorov. Teoretické štúdie a projekty sú odsimulované na rôznych typoch robotov od najjednoduchších až po tie najzložitejšie. Bez ohľadu na to, aké sú zložité a koľko stupňov voľnosti majú, každý takýto systém sa skladá z niekoľkých charakteristík: a) schopnosť otočenia na jednej nohe b) vzorkovanie chôdze, opakovanie rovnakých krokov symetrických c) správna zmena dvoch fáz, keď robot stojí na jednej nohe a keď robot stojí na dvoch nohách. Všetky kĺby robota sú ovládané motormi. Spojenie chodidla a zeme sa takisto dá považovať za kĺb, ktorý ale nie je ničím poháňaný. Tento kĺb je veľmi dôležitý pre stabilitu celého systému. Noha nemôže byť ovládaná priamo, a preto sa musí ovládať telo a ruky. Preto celkový indikátor správania sa mechanizmu je vlastne jeden bod, kde všetky sily pôsobiace na mechanizmus sú nahradené jednou silou. Tento bod je definovaný ako Zero Moment Point

(ZMP). Typy chôdze sa preto navrhujú v závislosti na tento bod.

## II. PRINCÍP METÓDY ZMP A PERSPEKTÍVY JEJ VYUŽITIA

V tejto časti uvedieme iba jadro princípu metódy ZMP. Podrobnejšie informácie o jej fungovaní je možné nájsť napr. v [2]. Sú dve základné možné interpretácie pojmu ZMP.

Interpretácia č.1.: ZMP je definované ako bod na zemi, v ktorom sieť momentov inerných síl a gravitačných síl nemá žiadnu zložku, silu pozdĺž horizontálnych osí.

Interpretácia č.2.: P je bod, kde  $T_x = 0$  a  $T_y = 0$ ,  $T_x$ ,  $T_y$  reprezentujú momenty okolo x-ovej a y-ovej osi generované reakčnou silou  $F_r$  a reakčným krútiacim momentom  $T_r$ . Bod P je definovaný ako ZMP. Ak tento bod je v podporovanom polygóne, systém je stabilný.

Častokrát je ZMP zle definovaný, napríklad že ZMP mimo podporovaného polygónu nemá zmysel, ale pravda je, že ZMP mimo podporovaného polygónu neexistuje, pretože hneď ako sa dostane mimo podporovaného polygónu, robot spadne na zem. Taktiež je veľmi dôležité ujasniť si pojmy center of pressure (CoP) ťažisko a ZMP. Tlak medzi chodidlom a zemou sa dá vždy nahradiť silou pôsobiaceou na ťažisko. Ak táto sila balansuje, vyrovnáva všetky aktívne sily pôsobiace na mechanizmus počas pohybu (inertné, gravitačné, momenty a centrifúgy), potom je to ZMP. Ak je teda krok dynamicky rovnovážny, ťažisko a ZMP sú v jednom bode. Ak krok nie je dynamicky rovnovážny, ZMP neexistuje a mechanizmus kolabuje.

Očakávaná od robotov stále rastú. Už dnes majú ľudia ambície použiť v robotoch komunikáciu s človekom za účelom zabrániť kritickým situáciám. Hlavné možnosti využitia ZMP v budúcnosti preto je možné vidieť v týchto úlohách:

- noha a chodidlo, ktoré sa viac približuje realistickej nohe kvôli zvýšeniu stability
- lepšia pružnosť v kĺboch, hlavne kolenných v prípade dopadu nohy na zem, čo je v niektorých materiáloch považované za doménu humanoidnej robotiky
- plynulejší prechod medzi jednotlivými fázami robota ako sú chôdza, chôdza po schodoch, vyhýbanie sa prekážkam alebo zrýchlenie chôdze
- špeciálna výzva je chôdza nezávislá na ruky, napríklad zobrať pohár zo stola pri chôdzi bez zastavenia
- rovnováha robota nielen na statických podlahách, ale aj na mobilných a dynamických systémoch s rôznymi stupňami voľnosti.

V doteraz pomenovaných systémoch existuje simulátor GHDS (General Human/Humanoid Dynamic Simulator). Jeho najväčším cieľom je dosiahnuť rovnaké výsledky v reálnom svete a simulátore. To znamená, že roboti budú musieť dosiahnuť také výsledky, aby boli pripravení aj na vysoké požiadavky reálnych aplikácií [2].

Výskumom špeciálnych typov humanoidných robotov sa zaoberajú napríklad aj tímy a výskumníci na celom svete, ktorí neustále vyvíjajú robota Nao, s ktorým sa prezentujú na celosvetových súťažiach v robotickom futbale v Standart Platform League Robo Cup. Jedným z týchto tímov je B-human, ktorý je v súčasnosti najúspešnejším tímom v robotickom futbale.

## III. TYPY KOPOV TÝMU B-HUMAN

B-human je názov pre nemecký tím vysokoškolských študentov, ale aj profesorov z Univerzity Bremen, fakulty matematiky a informatiky. Tento tím je výnimočný tým, že už niekoľko krát zvíťazil na majstrovstvách v robotickom futbale. Robotický futbal je veľmi podobný klasickému futbalu čo sa týka pravidiel, no je veľkosťou prispôsobený robotom Nao od firmy Aldebaran robotics. K tak veľkému úspechu ako napríklad víťazstvo 11:1 v jednom z ich zápasov prispieva hlavne veľmi dobre naprogramovaný celkový systém. Rozhodne k tomu prispieva aj ich systém kopu, stabilitu nevynímajúc.

Pri kope je dôležité udržať rovnováhu, nespádnúť. Každý kop je definovaný ako séria statických uhlov kĺbov. Každá séria je kľúčová. Všetky kĺby medzi dvoma kľúčovými sériami sú interpolované. Hlavná nevýhoda takéhoto prístupu je v jeho flexibilita. Keď raz pohyb začne, nedá sa ovplyvniť ani zastaviť, prijať nové informácie. Podobné prístupy sú vo viacerých robotických tímoch. B-human obohatil tento prístup o možnosť balansovania na kompenzovanie vonkajších otrasov. Kľúčové série uhlov kĺbov sú definované ako karteziánske pozície končatín, ktoré sú interpolované. Skutočné uhly kĺbov sú vypočítané pomocou inverznej kinematiky. Systém obsahuje modul na stabilizáciu, ktorý je založený na ťažisku a senzorech spätnej väzby uzavretého PID regulátora.

Pohybový modul je vyvinutý ako modul, ktorý môže byť použitý paralelne s ostatnými modulmi ako napríklad s modulom kráčania. Hlavná myšlienka je urobiť jednoduché krivky pohybov zložitejšími. Celý pohyb je potom rozdelený do takzvaných fáz, ako napríklad zdvihnutie nohy alebo kopnutie správnou nohou. Každá fáza má svoju dobu trvania a každá končatina obsahuje svoju krivku pohybu. Taktiež je nutné zabezpečiť, aby prechod medzi

fázami bol plynulý a hladký. Modul pohybu prijme zoznam pozícií všetkých kĺbov, ale aj napríklad informáciu o polohe lopty ako vstup. Ako výstup je samotný pohyb v závislosti na okolnosti.

Základná štruktúra modulu pohybu je zložená z niekoľkých krokov. Najskôr zistí ID pohybu, ktorý má byť vykonaný. Potom, na začiatku každej fázy sa aplikujú dynamické body na aktuálnu fázu a inicializujú sa trajektórie. Rovnako je zabezpečená aj plynulosť pohybu. Následne sa získajú pozície končatín v aktuálnom čase a aplikuje sa inverzná kinematika na výpočet uhlov kĺbov pre pohyb ostatných končatín na dosiahnutie stability, rovnováhy [4].

Princíp stability robota Nao v robotickom futbale je čiastočne porovnateľný, ako je riešený v tejto práci, aj keď v nej navrhované metódy fungujú na odlišnom princípe.

#### IV. NAVRHNUTÁ METÓDA

Udržanie rovnováhy humanoidného robota je pomerne náročný proces, ktorý sa samozrejme dá riešiť rôznymi spôsobmi. Implementovaný spôsob udržania rovnováhy humanoidného robota na jednej nohe je zložený z dvoch veľkých celkov a je pomerne jednoduchší ako stabilita riešená v tíme B-human. Riešenie stability robota Nao je navrhnuté za pomoci tlakových senzorov na spodku chodidla robota Nao a následným vykonaním pohybu stability za pomoci neurónových sietí.

Prvý celok je samotné vykonanie pohybu jednotlivých častí tela tak, aby robot Nao nespadol na zem, aby nestratil stabilitu, rovnováhu. To, aby udržal rovnováhu, môžu zabezpečiť napríklad ruky, noha, ktorá nedrží telo, ale je voľná vo vzduchu, samotný trup ale aj hlava. Program, ktorý bol navrhnutý, použil na korigovanie rovnováhy voľnú nohu a trup. Týmto dvoma časťami robot hýbe, aby udržal rovnováhu.

Druhý bod samotného programu je učenie. Učenie je veľmi dôležité z pohľadu správneho pohybu vyššie spomínaných častí tela na korigovanie rovnováhy. To znamená učenie robota Nao naučí, kedy, kde a o koľko sa majú jednotlivé časti tela pohnúť.

Tieto dva najväčšie celky programu sú úzko prepojené. Pri spustení programu robot Nao nevykonáva nič okrem toho, že sa postaví na jednu nohu. Najskôr sa v členku nakloní doľava a potom zdvihne nohu. Nie je možné urobiť tento postup naopak, lebo pri zdvihu nohy robot spadne doprava. Preto je potrebné najskôr v členku celého robota Nao nakloniť doľava, aby sa ťažisko robota premiestnilo trochu doľava a až potom následne zdvihnúť nohu. Takýto postup, algoritmus, môže byť pre estetickosť doplnený o nastavenie rúk do polohy lastovičky. To však nie je potrebné.

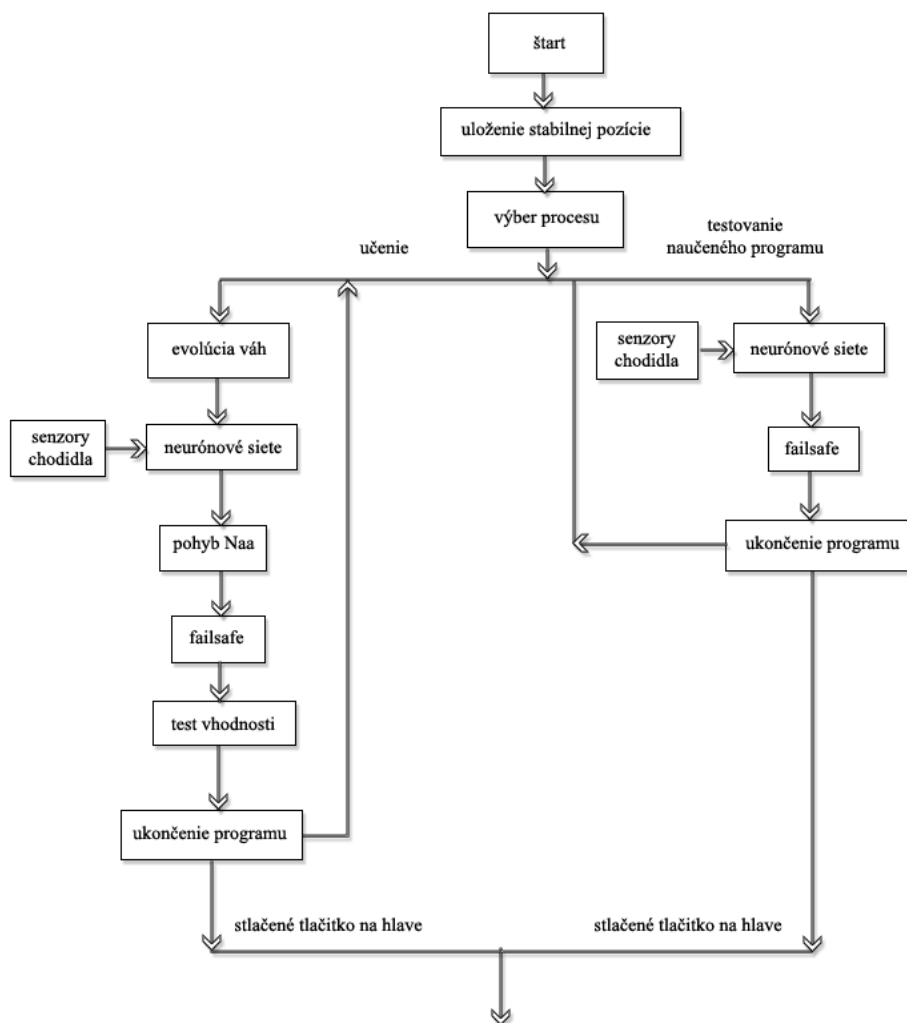
Po spustení a pohybe Naa do pozície postoja na jednej nohe robot čaká na používateľa. Tromi dotykovými tlačidlami na hlave sa vykonajú nami požadované operácie. Jedným tlačidlom sa vykoná operácia uloženia stabilnej pozície, to znamená uloženia senzorov z chodidla, ktoré sú v stabilnej polohe a druhým tlačidlom sa začne proces učenia. Tretie nám slúži na to, keď práve nechceme učiť Naa, ale iba vykonať stabilizáciu, pustiť Naa na ostro. Na obr. 1 je zobrazený hrubý návrh algoritmu učenia. Evolučný algoritmus vytvára váhy, ktoré sú použité v neurónových sieťach. Neurónové siete vypočítavajú aký pohyb má Nao urobiť.

#### V. IMPLEMENTÁCIA ALGORITMU

Ako implementačný algoritmus je použitý navrhnutý algoritmus. Kvôli B-human softwaru musel byť ale trochu pozmenený. Prvá zmena spočívala v tom, že robot pri stabilizácii nepoužíval nohu, ktorá je vo vzduchu a trup, ale obe horné končatiny. Konkrétne štyri časti horných končatín LShoulderPitch, RShoulderPitch, LShoulderRoll a RShoulderRoll. Namiesto dvoch boli použité štyri neurónové siete. Nohou, ktorú mal Nao vo vzduchu, nemohol zabezpečovať rovnováhu, keďže ňou kope do lopty. Trup priamo súvisí s oboma nohami a nie je možné hýbať trupom pri kope do lopty.

Ďalšou úpravou bolo, že proces učenia vznikol na navrhnutom systéme a do implementovaného algoritmu sa vložila časť už naučeného systému.

Nahradená bola funkcia `simpleCalcArmJoints()` novou metódou. Funkcia `simpleCalcArmJoints()` sa nachádza vo funkcii `calcJoints()` a stabilitu Naa rieši pomocou inverznej kinematiky. Táto funkcia je pri B-human volaná dvakrát pre ľavú a dvakrát pre pravú ruku.



Obr. 1 Algoritmus navrhnutej metódy.

Algoritmus implementovanej metódy pracuje najskôr tak, že zistí hodnoty na všetkých štyroch tlakových senzoch chodidla nohy, ktorá stojí pevne na zemi. Tieto hodnoty sú zapísané do pomocných premenných. Pomocou štyroch neurónových sietí sa vypočítajú hodnoty, ktoré sa následne použijú na pohyb horných končatín. Pre správne vloženie hodnôt do neurónových sietí bolo potrebné tieto zistené hodnoty najskôr predeliť číslom 4000 z dôvodu použitia sigmoidálnej aktivačnej funkcie, pretože hodnoty z tlakových sensorov dosahujú rozsah od 0 po 4. Pre silnejší pohyb boli potom výsledné hodnoty vynásobené číslom 100. Konkrétne riešenie pre pohyb horných končatín spočíva v zistení aktuálnej pozície jednotlivého kĺbu a pripočítanie výslednej hodnoty vynásobenej číslom 100. Čísla 4000 a 100 po niekoľkých opakovaných pokusoch vyšli ako najúčinnějšíe.

## VI. VÝSLEDKY EXPERIMENTOV

Výsledky implementovanej metódy sú zaznamenané v tabuľkách. Sila kopu je udaná v milisekundách, keďže sa jedná o trvanie švihy nohy.

Vzdialenosť je dĺžka od robota k dopadu vykopnutej lopty. Je udávaná v centimetroch.

Ďalšie veličiny sa týkajú hodnôt nameraných uhlov horných končatín. Celkový súčet týchto hodnôt udáva pozíciu rúk k telu. Čím je suma týchto hodnôt menšia, tým sú horné končatiny robota viac pri tele a tým pádom je robot ustálený a pripravený na ďalšiu fázu pri robo-futbale. Ak je sila kopnutia príliš veľká, dôjde k jemnému otočeniu robota. Táto hodnota je udávaná v stupňoch. Čím menšie je natočenie, tým je robot stabilnejší a jeho poloha sa pri futbale nezmení. V prípade veľkej sily kopnutia by mohlo dôjsť aj k strate rovnováhy a robot by mohol spadnúť. Pri meraní so silou 200 a 100 ms neboli namerané konkrétne hodnoty, nakoľko veľké ihrisko je v porovnaní s reálnym futbalovým ihriskom oveľa menšie. Preto boli hodnoty vyššie ako 250 cm zaokrúhlené na hodnotu 250 cm.

Tabuľka 1  
Výsledky implementovanej metódy

	Implementovaná metóda				Priemer
Sila (dĺžka v ms)	500	400	300	200	
Vzdialenosť (cm)	69,8	111,6	162,8	250	148,55
Súčet uhlov kĺbov	409,56	433,94	446,62	462,98	438,275
Natočenie (stupne)	0	0	3,6	9,2	3,2

Pri metóde B-human bol zvolený rovnaký postup experimentov ako aj zobrazenia výsledkov.

Tabuľka 2  
Výsledky B-human metódy

	B-human metóda				Priemer
Sila (dĺžka v ms)	500	400	300	200	
Vzdialenosť (cm)	49,6	82,4	178,2	244,4	138,65
Súčet uhlov kĺbov	611,9	612,72	610,74	611,3	611,665
Natočenie (stupne)	0	0	3	9,2	3,05

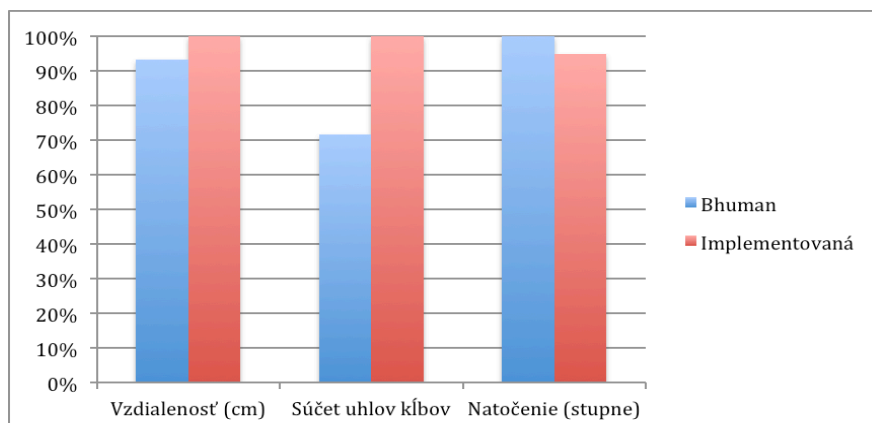
## VII. POROVNANIE VÝSLEDKOV OBOCH METÓD A ZÁVER

V grafe na obr. 2 je vyobrazené percentuálne porovnanie jednotlivých prvkov merania. Grafy sú rozdelené podľa sily kopy. Z výsledkov vyplýva, že pri sile 500 ms je implementovaná metóda lepšia z hľadiska vzdialenosti dokopnutia a pozície horných končatín. V tomto prípade sa robot nenatočil vôbec, z čoho vyplýva, že bol stabilný. Veľmi podobné výsledky sú dosiahnuté aj pri sile 400 ms. Pri sile 300 ms výsledky meraní boli lepšie u B-human, a to vo vzdialenosti dokopnutia a natočenia robota. Napriek lepším výsledkom u B-human pri natočení robota sú však tieto hodnoty pomerne rovnaké, keďže natočenie bolo v priemere trojstupňové. Pri sile 200 ms sú výsledky implementovanej metódy lepšie ako metódy B-human. Merania pri sile 100 ms nie sú zaznamenané, nakoľko boli skoro identické s hodnotami pri sile 200 ms.

Celkovo ale je možné zhodnotiť, že implementovaná metóda stability robota Nao pri kope je lepšia s porovnaním metódy B-human. Napriek tomu ale robot nestratil rovnováhu ani pri jednej z experimentálnych metód. Z toho vyplýva, že obe metódy sú pomerne stabilné.

Tabuľka 3  
Porovnanie oboch metód

	B-human metóda	Implementovaná metóda	percentuálny rozdiel
Vzdialenosť (cm)	138,65	148,55	6,66
Súčet uhlov kĺbov	611,665	438,275	39,56
Natočenie (stupne)	3,05	3,2	-4,92



Obr. 2 Grafické porovnanie výsledkov oboch metód.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] H. Wongsuwarn, and D. Laowattana, "Neuro-Fuzzy Algorithm for a Biped Robotic System," In: *Proceedings of World Academy of Science, Engineering and Technology*, vol. 2, 2008, pp. 740-746.
- [2] M. Vukobratovic and B. Borovac, "Zero-moment point - Thirty five years of its life," *International Journal of Humanoid Robotics*, vol. 1, 2004, pp. 157-173.
- [3] M. Sciranka, "Lokomočné mechanizmy kráčajúcich robotov," In: *1st Annual International Ph.D. Conference AVS/PMA*, Praha, pp. 134-138.
- [4] M. LaValle, *Planning Algorithms*. Cambridge, U.K.: Cambridge University Press, 2006.

# Fuzzy logika pri detekcii poisťovacích podvodov

Ján VAŠČÁK

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

jan.vascak@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá prehľadom aplikácií fuzzy logiky v jednotlivých poisťovacích segmentoch, popisom problematiky poisťných podvodov a spôsobmi ich odhaľovania. Následne zahŕňa návrh a implementáciu fuzzy modelu pre podporu odhaľovania podvodných poisťných pohľadávok, pričom implementácia riešenia obsahuje okrem samotného vytvorenia a testovania funkčného modelu aj programové upravenie využívajúceho softvéru ako aj vyhodnotenie niektorých druhov experimentov.

**Kľúčové slová** — fuzzy logika, fuzzy expertný systém, poisťovníctvo, odhaľovanie poisťných podvodov

## I. ÚVOD

Účelom príspevku je podať aktuálny obraz o využití fuzzy logiky v poisťovníctve – uviesť príklady aplikácii fuzzy expertného systému v poisťovacej doméne a ozrejmiť problematiku poisťovacích podvodov. Následne bolo hlavným cieľom tohto príspevku návrh a implementácia fuzzy modelu, ktorý je schopný na základe súboru vstupných parametrov týkajúcich sa poisťníka vyhodnotiť mieru pravdepodobnosti, akou je poisťná udalosť podvodná. Tento model by tak mohol slúžiť ako forma podpory, pomocou ktorej, by poisťovníka vedela rozhodnúť o podvodnosti poisťnej pohľadávky. Predkladaný model by však na rozdiel od mnohých iných existujúcich riešení pracoval s fuzzy logikou a nebol by teda striktno obmedzovaný limitami booleovskej logiky - dovoľoval by teda vykonávať operácie s vágnymi výrazmi a tým pádom viachodnotovým opodstatňovaním.

## II. POISŤNÉ PODVODY A ICH ODHAĽOVANIE

Pod poisťným podvodom rozumieme taký čin, ktorý je spáchaný s úmyslom podvodne získať nejaký prospech či výhodu, ktorý inak nie je poisťníkovi pridelený [11]. Poisťný podvod môže byť spáchaný dvoma hlavnými spôsobmi: buď sfaľšovaním poisťnej udalosti a prezentovaním tejto ako poisťnej udalosti s nárokom na vyplatenie poisťného plnenia, alebo zamľčaním dôležitých faktov, resp. zavádzaním poisťovateľa pri procese prijímania poisťného rizika s úmyslom získania výhodnejšej poisťnej sadzby či poisťného poplatku. Spáchanie poisťného podvodu sa považuje za trestný čin, ktorého sa dopustí jednak samotný poisťník, keď mu je neprávom poskytnuté poisťné plnenie, no rovnako sa za trestný čin považuje už aj samotný pokus o spáchanie podvodu, či spolupráca na ňom. Hlavným motívom pre konanie týchto podvodov je finančný zisk vychádzajúci s viazanej povinnosti poisťovateľa uhradiť poisťníkovi poisťné plnenie [11].

Poisťné podvody predstavujú pre poisťovne vážny problém, keďže majú za následok veľké finančné straty. Preto sú poisťovne nútené reagovať na tento problém a takéto falošné poisťné udalosti vedieť odhaliť. Podľa viacerých štatistických odhadov sa v súčasnosti miera pochybných poisťných udalostí a podvodov pohybuje v priemere medzi 10% až 15% všetkých poisťných udalostí, pričom do tohto čísla sa rátaajú iba tie podvody, ktoré sa podarilo objasniť. Úspešnosť odhaľovania poisťných podvodov je vôbec jedným z najdôležitejších faktorov, ktoré ovplyvňujú rentabilitu poisťovní [9].

Kvôli veľmi širokému spektru existujúcich poisťných podvodov sa budeme ďalej orientovať na poisťné podvody týkajúce sa poistenie motorových vozidiel – podvody fyzického poškodenia automobilov. Tento druh podvodu je známy aj pod skratkou APD (*auto physical damage*). Ide



o jeden z najrozšírenejších druhov poistných podvodov, keď poisťník skreslí skutočnosť, resp. poskytne poisťovateľovi falošné informácie o vyskytnutej poistnej udalosti na základe ktorej sa od poisťovateľa dožaduje poistného plnenia. V prípade tohto APD podvodu poisťník klame o fyzickom poškodení motorového vozidla ku ktorému došlo. Toto fyzické poškodenie zahŕňa udalosti ako: kolíziu alebo uhorenie vozidla, krádež vozidla, vandalizmus či zničenie následkom živeľnej udalosti (napr. krupobitím) [7].

Veľmi častým, ak nie vôbec najčastejším typom klamania poisťovateľa ohľadom fyzického poškodenia auta je predstieranie krádeže automobilu. Ide o podvod, ktorý je veľmi ťažko dokázať. Typickým príkladom je to, keď poisťník tajne predá svoj automobil tretej strane, resp. ho rozoberie a predá na diely a následne z dôvodu krádeže od poisťovateľa žiada vyplatenie poistného plnenia [8]. Inscenácia krádeže auta má taktiež za úlohu zbavenie sa zodpovednosti za vzniknutú škodu v prípade dopravnej nehody alebo zakrytím technickej poruchy, či poškodeniu ktoré poisťník pripíše zlodejovi. Dokazovanie takejto podvodnej pohľadávky je komplikované najmä kvôli tomu, že na jednoznačné potvrdenie či vyvrátenie sú nutné servisné záznamy, ktoré často chýbajú, alebo sú taktiež zmanipulované.

Ďalej, nasleduje prehľad indikátorov a faktorom podstatných pri procese odhaľovania APD podvodov. Kvôli rozsahu popisu a množstva týchto indikátorov nie je možné uviesť všetky.

Príklady faktorov dôležitých pri odhaľovaní:

- Miesto z ktorého bol automobil ukradnutý

Tento faktor má za úlohu preukázať vierohodnosť krádeže. Ak klient nahlásil miesto krádeže, ktoré má veľmi malú, alebo skoro žiadnu mieru výskytov krádeží, je zrejmé, že automobil na tomto mieste ukradnutý nebola klient pravdepodobne klame. Naopak, miesta ako parkoviská pri

obchodných centrách, alebo veľké verejné parkoviská majú veľkú mieru krádeží.

- Osoba, ktorá ukradnuté vozidlo našla

Pokiaľ odcudzené motorové vozidlo vypátra samotný poisťník, alebo jeho príbuzný, či známy, je veľmi vysoká pravdepodobnosť, že krádež vozidla bola iba nafingovaná. V tomto smere je nutné vždy zistiť informácie o osobe, ktorá ukradnuté vozidlo vypátrala a jej vzťahu k poisťníkovi.

Ďalej je tu definovaný pojem *model pre odhaľovanie poistných podvodov* ako nástroj, prostredníctvom ktorého sú poisťovne schopné predchádzať stratám odhaľovaním pochybných poistných pohľadávok. Tieto modely sú založené na rozličných prístupoch a platformách. Hlavným cieľom je vytvorenie takéhoto modelu na podporu odhaľovania podvodných poistných pohľadávok, ktorý by ale bol založený na fuzzy logike.

### III. NÁVRH A IMPLEMENTÁCIA FUZZY MODELU PRE PODPORU ODHAĽOVANIA PODVODNÝCH POISTNÝCH POHLADÁVOK

Cieľom praktickej časti tohto návrhu bolo vytvorenie funkčného modelu pre odhaľovanie poistných podvodov založeného na fuzzy logike. Takýto model vo forme fuzzy expertného systému by dokázal na základe súboru vstupných parametrov určiť ako výstupnú hodnotu index podvodnosti – mieru určujúcu pravdepodobnosť toho, že je daná poistná pohľadávka podvodná. Tento index bude určovať výstupné hodnoty v intervale 0 až 1 a bude ho teda možno vyjadriť ako percentuálnu pravdepodobnosť podvodnosti prípadu.

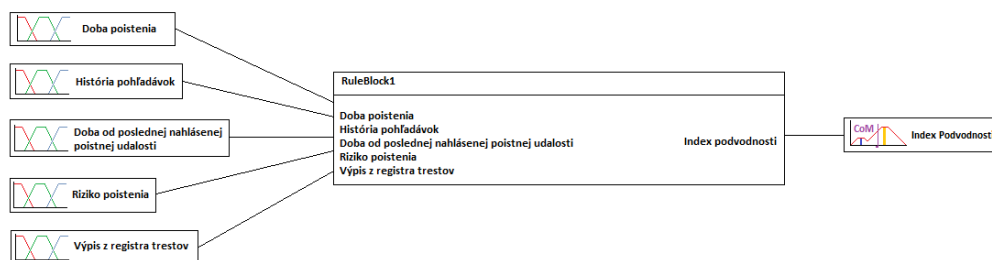
Vybranou poisťovacou doménou pre tento model bude oblasť poistenia motorových vozidiel, no model sa nebude zameriavať ani tak na konkrétne indikátory súvisiace s poistnými udalosťami, ale skôr na poisťníkovu reputáciu a jeho dôveryhodnosť. Predpokladáme totiž, že existuje priamy vzťah medzi hodnovernosťou poisťníka a podvodnosťou pohľadávky. Na základe vyhodnoteného poisťníkovho renomé teda systém vyhodnotí pravdepodobnosť toho, že poistná udalosť, ktorú nahlásuje, je podvodná. Na realizáciu návrhu bol využitý open-source softvér jFuzzyLogic, napísaný na platforme Java. Tento balík implementuje FuzzyControlLanguage (FCL) a FuzzyInferenceSystem (FIS), prostredníctvom ktorých je možné zjednodušene definovať vlastné lingvistické premenné, fuzzy množiny, funkcie príslušnosti či fuzzy pravidlá a vykonávať experimenty [3].

Znalostná báza fuzzy expertného systému pre podporu odhaľovania odvodných poistných pohľadávok bude obsahovať päťicu nasledovných vstupných premenných:

- Doba poistenia;
- História pohľadávok;
- Doba od poslednej nahlásenej poistnej udalosti;
- Riziko poistenia;

- Výpis z registra trestov;

Tieto vstupné parametre sa po načítaní ich hodnoty fuzzifikujú a prostredníctvom tradičného (jednoiteračného) spôsobu usudzovania sa vyhodnotí výstupná hodnota – index podvodnosti.



Obr. 1. Štruktúra usudzovania.

#### IV. EXPERIMENTY A ICH VYHODNOTENIE

Na testovanie implementovaného modelu sme vygenerovali vzorku desiatich poistných udalostí, pričom každá poistná udalosť je jedinečná a jednotlivé poistné udalosti popisujú rozličné možné situácie. Vzorku sme generovali pretože sa nám nepodarilo z dôvodu citlivosti údajov získať reálne a zároveň relevantné informácie od poisťovní. Pri popise experimentov sme uvádzali

:

- Popis situácie
- Hodnoty vstupných premenných
- Aktivované fuzzy pravidlá
- Grafické znázornenie inferencie
- Hodnotu výstupnej premennej
- Zhodnotenie výsledku experimentu

Experimenty boli vykonávané pri týchto podmienkach:

*Metóda defuzzifikácie:* Center of Gravity (metóda ťažiska)

*AND metóda:* metóda MIN;

*Aktivačná metóda:* metóda MIN;

*Akumulačná metóda:* metóda MAX;

V nasledujúcom zo súboru vykonaných experimentov ako ukážku uvádzame dva z nich.

##### A. Experiment 1

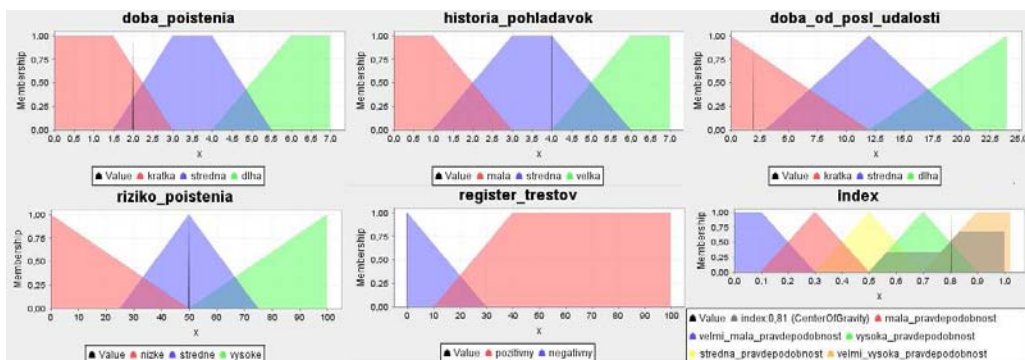
*Popis situácie:* Poistník nahlásil poisťovní počas svojho iba dvojročného poistenia už štvrtú poistnú udalosť, pričom jeho posledná poistná udalosť bola hlásená iba pred dvoma mesiacmi a jeho poistné riziko je 50%.

*Hodnoty vstupných premenných:* Doba poistenia – 2 roky, História pohľadávok – 4 pohľadávky, Posledná nahlásená poistná udalosť – 2 mesiace, Riziko poistenia – 50%, Výpis z registra trestov – 0%.

*Aktivované fuzzy pravidlá:* RULE 17 : IF doba\_poistenia IS kratka AND historia\_pohladavok IS stredna AND doba\_od\_posl\_udalosti IS kratka AND riziko\_poistenia IS stredne THEN index IS velmi\_vysoka\_pravdepodobnost. Grafické znázornenie inferencie je na obr. 2.

*Hodnota výstupnej premennej:* Index podvodnosti = 0.8058

*Zhodnotenie výsledku experimentu:* Táto poistná pohľadávka je z veľmi vysokou pravdepodobnosťou (80.58%) podvodná.



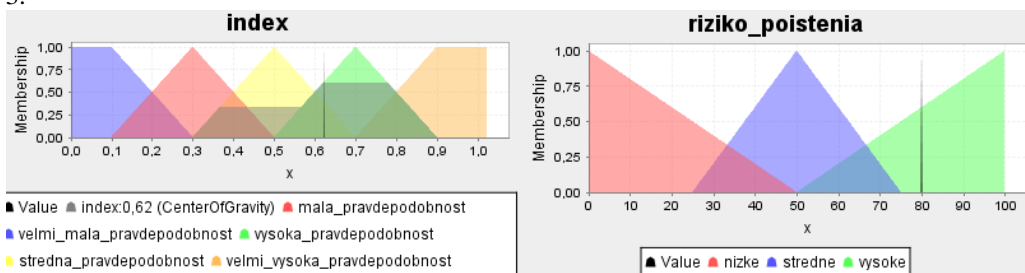
Obr. 2. Grafická inferencia pre Experiment č 1.

### B. Experiment 2

*Popis situácie:* Rovnaká situácia ako v experimente 1, avšak riziko poistenie je 80%.

*Hodnoty vstupných premenných:* Doba poistenia – 2 roky, História pohľadávok – 4 pohľadávky, Posledná nahlásená poistná udalosť – 2 mesiace, Riziko poistenia – 80%, Výpis z registra trestov – 0%.

*Aktivované fuzzy pravidlá:* RULE 18 : IF doba\_poistenia IS kratka AND historia\_pohladavok IS stredna AND doba\_od\_posl\_udalosti IS kratka AND riziko\_poistenia IS vysoke THEN index IS vysoka\_pravdepodobnost. Grafické znázornenie inferencie je na obr. 3.



Obr. 3. Grafická inferencia pre Experiment č 2.

*Hodnota výstupnej premennej:* Index podvodnosti = 0.6242

*Zhodnotenie výsledku experimentu:* Zvýšené poistné riziko zmiernilo pravdepodobnosť podvodnosti o takmer 20%, takže pravdepodobnosť, že poistná udalosť je podvodom je 62.42%. Celkové vyhodnotenie všetkých experimentov je v tabuľke 1.

Tabuľka 1  
Celkové vyhodnotenie všetkých experimentov.

	Doba poistenia (roky)	História pohľadávok (pohľadávky)	Doba od posledného nahlásenia poistnej udalosti (mesiace)	Riziko poistenia (%)	Výpis z registra trestov (%)	Index podvodnosti
Experiment 1	2	4	2	50	0	0.8058
Experiment 2	2	4	2	80	0	0.6242
Experiment 3	6	1	18	25	0	0.1264
Experiment 4	6	1	18	25	60	0.5963
Experiment 5	3.5	2	12	30	0	0.4000
Experiment 6	4	7	3	85	15	0.7231
Experiment 7	4	7	3	50	70	0.8913
Experiment 8	1.75	1	8	20	0	0.5319
Experiment 9	5.25	4	1	70	0	0.4356
Experiment 10	5.25	4	24	70	0	0.3585

## V. ZÁVER

Zámerom tohto príspevku bolo vytvoriť prehľad teoretických východísk fuzzy logiky v oblasti poisťovníctva a poistných podvodov a predložiť vlastné riešenie problému ich odhaľovania. Toto praktické riešenie sa od ostatných modelov pre účel odhaľovania poistných podvodov líši tým, že neoperuje ani tak z konkrétnymi indikátormi podvodu, ako skôr s činiteľmi, prostredníctvom ktorých systém dokáže odhadnúť mieru hodnovernosti poistníka.

Tento prístup ponúka na jednej strane možnosť účinne analyzovať poisníckove renomé, no na druhej strane neberie do úvahy konkrétne indikátory, ktorú súvisia bezprostredne s poisťnou udalosťou.

Prvá časť príspevku pojednáva o teórii fuzzy logiky a fuzzy expertných systémov v poisťovníctve ako takom a následne v jednotlivých oblastiach poisťovníctva akými sú klasifikácia a prijímanie poisťného rizika, či tvorba poisťných sadziieb a cien. Druhá časť príspevku sa venovala už priamo problematike poisťných podvodov a ich odhaľovania. V tejto kapitole sme sa z dôvodu mimoriadne širokej domény zamerali hlavne na podvody poistenia automobilov. Rozobrali sme v nej hlavné indikátory odhaľovania podvodných udalostí a predložili prehľad metód a prístupov ich vyšetrovania. Tretia časť príspevku sa zaoberala návrhom a implementáciou fuzzy modelu pre podporu odhaľovania podvodných poisťných pohľadávok. Táto časť sa skladala z programovej modifikácie využívajúceho softvéru jFuzzyLogic a implementácie navrhnutej bázy znalostí. V poslednej, štvrtej časti z názvom Experimenty sme vykonali testovanie navrhnutého a implementovaného modelu. Na vzorke desiatich rozličných poisťných udalostí sme preukázali funkčnosť a užitočnosť vytvoreného fuzzy expertného systému.

#### POĎAKOVANIE

Výskum podporovaný Národným projektovým grantom pre výskum a vývoj 1/0667/12 „Inkrementálne metódy učenia pre inteligentné systémy“ 2012-2015.

#### LITERATÚRA

- [1] K. Aho, "Benchmarking main activation functions in fuzzy cognitive maps," *Expert Systems Applications*, vol. 36, no. 3, 2009, pp. 5221–5229.
- [2] E. B. Belhadji, G. Dionne, and F. Tarkhani, "The design and analysis of an efficient local algorithm for coverage and exploration based on sensor network deployment," *IEEE Transactions on Robotics*, vol. 23, no. 4, 2007, pp. 661–675.
- [3] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, "KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 17, no. 2–3, 2011, pp. 255–287.
- [4] M. Grega, *Prehľad využitia fuzzy logiky v ekonomických aplikáciách*, bakalárska práca, Katedra kybernetiky a umelej inteligencie, Technická univerzita v Košiciach, 2011.
- [5] E. Markar, "Fuzziness in the semantic web: Survey and future directions," in *Proc. SEKE'08*, 2008, pp. 643–648.
- [6] A. Shapiro, "Integrating design stages of fuzzy systems using genetic algorithms," in *Proc. The second IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE)*, San Francisco, USA, 1993, pp. 413–417.
- [7] A. Ward, "Comparison of a self-organizing migration algorithm with simulated annealing and differential evolution for automated waveform tuning," *Advances in Engineering Software*, vol. 36, no. 10, 2005, pp. 645–653.
- [8] J. H. Wilson, *A learning system based on genetic adaptive algorithms*, Ph.D. dissertation, Department of Computer Science, University of Pittsburgh, USA, 1980.
- [9] R. Axelrod, *Structure of Decision: The Cognitive Maps of Political Elites*. Princeton University Press, 1976.
- [10] M. Gavalec and K. Mls, "Trend evaluation in on-line decision making," in *Proc. KOI 2008*, Pula, Croatia, ser. International Conference on Operational Research, 2008, pp. 267–274.
- [11] S. M. LaValle, *Planning Algorithms*. Cambridge, U.K.: Cambridge University Press, 2006.
- [12] L. E. Parker, "Evaluating success in insurance fraud detection: Experiences from ALLIANCE architecture implementations," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 13, no. 2, 2001, pp. 95–98.
- [13] R. Prado, S. García-Galán, J. Muñoz Expósito, and A. Yuste, "Knowledge acquisition in fuzzy-rule-based systems with particle-swarm optimization," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 6, 2010, pp. 1083–1097.
- [14] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufman Publishers, 1993.

# Fuzzy kognitívne mapy v robotických úlohách kopu

Ján Vaščák

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika

jan.vascak@tuke.sk

**Abstrakt** – Táto práca popisuje tvorbu dynamického robotického kopu, ktorý je založený na aproximácií Bézierových kriviek pomocou fuzzy kognitívnych máp. Predstavené sú rôzne varianty fuzzy kognitívnych máp. Príčinné väzby fuzzy kognitívnych máp sú určované pomocou algoritmu particle swarm optimization. Pre pokusy je použitý humanoidný robot Nao s využitím rozhrania hráčskeho systému robotického futbalu B-Human.

**Kľúčové slová** – fuzzy kognitívna mapa, optimalizácia, roj, kop, Bézierová krivka, Nao robot

## I. ÚVOD

Veľká pozornosť pre úspešné hranie futbalu s humanoidnými robotmi, je venovaná nielen vývoju rýchlej ale aj robustnej chôdze. Napríklad môže to byť flexibilita, pokiaľ ide o náhle zmeny smeru chôdze, aby bolo možné kompenzovať poruchy, ako je kolízia s inými robotmi alebo nerovnosť povrchu. Tieto dva druhy robustnosti sú predpokladom pre druhý najdôležitejší druh pohybu, kopu.

V [1] je popísaný obrazový prístup, ktorý je jeden z najbežnejších prístupov pri projektovaní pohybu. Daná metóda definuje pohyby ako sadu sérií statických uhlov kĺbu.

V [2] je obrazový prístup rozšírený o možnosť reagovať na vonkajšie zmeny prostredia. V danom systéme, sú kľúčové obrazy definované pomocou karteziánskych pozícií končatín, ktoré sú následne interpolované. Spoločné uhly sú vypočítavané podľa inverznej kinematiky a Bézierových kriviek.

Nevýhodou oboch prístupov je, že sa berie do úvahy len poloha lopty, podľa ktorej je kop vygenerovaný.

Prínosom tejto práce je nový prístup pre generovanie kopu. Metóda generuje kop na základe polohy lopty a na základe iných informácií získaných z prostredia (napr.: poloha brány, poloha prekážok), ktoré slúžia ako vstupy do fuzzy kognitívnej mapy. Na výstupe fuzzy kognitívnej mapy budú riadiace body pre inverznú kinematiku.

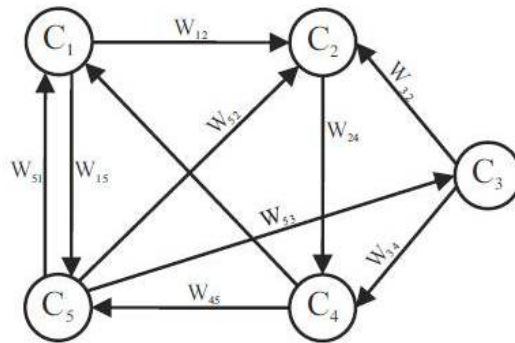
## II. FUZZY KOGNITÍVNE MAPY A ICH UČENIE

### A. Fuzzy kognitívne mapy

Fuzzy kognitívne mapy (FKM) sú metódy modelovania, ktoré kombinujú princípy fuzzy logiky a neuronových sietí. Boli vyvinuté Koskom ako rozšírenia kognitívnych máp a patria do skupiny neuro-fuzzy systémov, ktoré sú schopné začleniť ľudské vedomosti a prispôbiť ich prostredníctvom vzdelávacích postupov [3].

FKM majú širokú oblasť použitia, vrátane modelovania komplexných a inteligentných systémov, analýzu rozhodovania, a grafovú analýzu správania [4]. Taktiež boli použité pre plánovanie a rozhodovanie v oblasti medzinárodných vzťahov a sociálnych systémov, rovnako ako v manažovaní vedy a výskumu [5]. Dickerson a Kosko používali FKM na simuláciu virtuálnych svetov [6]. Okrem toho boli FKM používané pre modelovanie systémov videnia, v oblasti plánovania systémov radiačnej terapie, ale aj pre triedenie nádorov močového mechúra pomocou charakteristík ochorenia [7].

FKM sa skladá z konceptov,  $C_i$ ,  $i = 1, \dots, N$ , kde  $N$  je celkový počet konceptov. Každý koncept predstavuje jeden z kľúčových faktorov daného systému a je charakterizovaný hodnotou,  $A_i \in [0, 1]$ ,  $i = 1, \dots, N$ . Koncepty sú vzájomne prepojené pomocou váhovej hodnoty, ktorá naznačuje vzťahy medzi konceptmi. Jednoduchá FKM s piatimi konceptmi a desiatimi váhovými



Obr. 1 Fuzzy kognitívna mapa s piatimi konceptmi  $C$  a desiatimi váhovými hodnotami  $W$ .

hodnotami je znázornená na obr. 1. Každé prepojenie medzi dvoma konceptmi,  $C_i$  a  $C_j$ , má váhu,  $W_{ij}$ , ktorá je analogická k príčinnej súvislosti medzi  $C_i$  a  $C_j$  [3].

Váha  $W_{ij}$  udáva, či je vzťah medzi danými dvoma konceptmi priamy alebo inverzný. Smer kauzality udáva, či koncept  $C_i$  spôsobí na koncept  $C_j$  alebo naopak.

Existujú tri typy váh:

- $W_{ij} > 0$ , vyjadruje pozitívnu príčinnosť.
- $W_{ij} < 0$ , vyjadruje negatívnu príčinnosť.
- $W_{ij} = 0$ , medzi konceptmi neexistuje príčinnosť.

Ľudské vedomosti a skúsenosti v systéme určujú typ a počet uzlov, rovnako ako aj váhy FKM. Skúsenosti a vedomosti sú vyjadrené pomocou váhovej matice, ktorá vyzerá nasledovne:

$$\mathbf{W} = \begin{pmatrix} W_{11} & W_{12} & \dots & W_{1N} \\ W_{21} & W_{22} & \dots & W_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1} & W_{N2} & \dots & W_{NN} \end{pmatrix} \quad (1)$$

Hodnota  $A_i$  konceptu  $C_i$ , vyjadruje množstvo zodpovedajúcej fyzikálnej veličiny a je odvodená od transformácie fuzzy hodnoty, ktorá je priradená odborníkmi alebo evolučnými výpočtami na číselné hodnoty. Po priradení hodnôt na koncepty a váhy, FKM konverguje k ustálenému stavu cez proces interakcie. V každom kroku je hodnota konceptu  $A_i$  zmenená hodnotami váhy konceptu, ktorý je k nemu pripojený a FKM je aktualizovaná podľa nasledujúceho vzorca [3],

$$A_i(k+1) = f \left( A_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^n W_{ij} A_j(k) \right) \quad (2)$$

kde  $k$  značí čítač iterácií a  $W_{ji}$  je váhová matica. Funkcia  $f$  je sigmoidálna funkcia,

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \quad (3)$$

kde  $\lambda > 0$  je parameter, ktorý určuje strmlosť funkcie v oblasti okolo nuly.

Daná funkcia je zvolená, pretože hodnoty konceptov  $A_i$ , musia ležať v intervale  $[0, 1]$ . Interakciami FKM je možné dosiahnuť rovnovážny stav po niekoľkých iteráciách, tj. hodnoty konceptov sa ďalej nemenia. Požadované hodnoty na výstupe konceptov zaručujú riadne fungovanie simulovaného systému.

### B. Návrh príčinných väzieb pomocou evolučných výpočtových techník

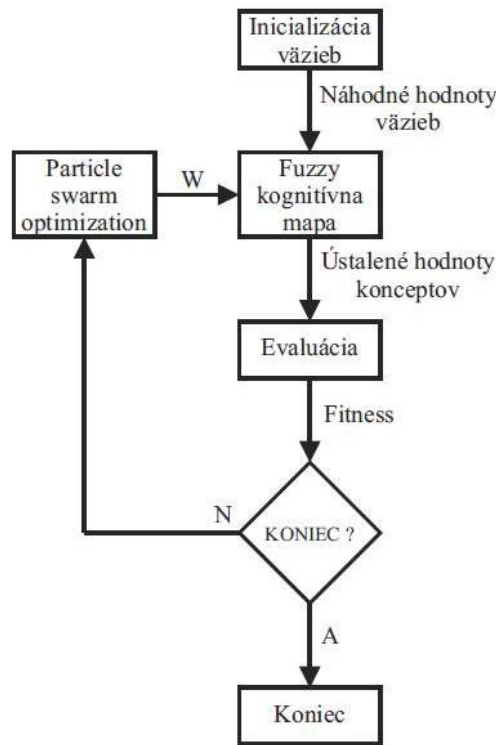
Podstatou tejto triedy algoritmov je pohyb v priestore prehľadávania. Keďže v prípade týchto algoritmov každý bod tohto priestoru zastupuje jedno kompletné riešenie (presnejšie povedané kandidáta na riešenie, pričom rôzni kandidáti v princípe predstavujú rôzne kvalitné riešenia, od optimálneho riešenia cez vyhovujúce až k neprijateľným), tak tento priestor býva označovaný aj ako priestor (kandidátov) riešení.

Prehľadávanie tohto priestoru v sebe typicky kombinuje generovanie resp. výber kandidátov na riešenie s testovaním týchto kandidátov. Pri riešení praktických úloh je potrebné preskúmať veľké množstvo kandidátov, kým je nájdené uspokojivé riešenie. Samotný pohyb v priestore prehľadávania je špecifický pre ten ktorý prehľadávací algoritmus [8].

### III. NAVRHOVANÝ PRÍSTUP

#### A. Algoritmus učenia FKM

Na začiatku činnosti algoritmu sa náhodne generujú príčinné väzby v intervaloch, ktoré určil odborník alebo laik. Taktiež je potrebné určiť, ktoré koncepty sú vstupné a ktoré výstupné. Ďalším krokom je evaluácia FKM. Následne sa vyhodnotia ustálené hodnoty konceptov, určí sa fitness vygenerovaných väzieb. Vzhľadom na fitness, algoritmus particle swarm optimization (PSO) určí zmenu príčinných väzieb. Ako ukončovacia podmienka slúži počet iterácií. Činnosť algoritmu je zobrazená na obr. 2.



Obr. 2 Priebeh algoritmu učenia FKM. W predstavuje upravené hodnoty príčinných väzieb.

Existuje mnoho váhových matíc, ktoré vedú FKM ku konvergencii do požadovaných oblasti. PSO je stochastický algoritmus, a preto je celkom prirodzené, že sú vygenerované rôzne matice pri rovnakom pokuse. Všetky výsledné matice sú vhodné pre konštrukciu FKM a následné obmedzenie daného problému, aj keď môže mať každá matica odlišný fyzikálny význam pre systém.

Vyššie uvedený algoritmus využíva iba primitívne informácie od odborníkov. Akékoľvek iné informácie, ktoré sú k dispozícii, môžu byť začlenené k zvýšeniu efektivity riadenia, a to buď tým, že sa zmení účelová funkcia s cieľom využiť dostupné informácie, alebo zavedením ďalšieho obmedzenia váh.

Jedinec je reprezentovaný pomocou váhovej matice  $W$ , pričom vzhľadom na algoritmus jedinec má znalosť o svojej najlepšej pozícii ( $LBest$ ) o najlepšej pozícii vzhľadom na roj ( $GBest$ ).

Znalosti o pozíciách sú reprezentované vektorom, kde jednotlivé časti reprezentujú atribúty v priestore prehľadávania.

Váhová matica  $W$  je reprezentovaná vektorom, ktorý sa skladá v poradí z jednotlivých váh, s výnimkou prvkov hlavnej diagonály,  $W_{11}, W_{22}, \dots, W_{nn}$ , ktoré sú podľa definície rovné nule.

$$X = \underbrace{[W_{12}, \dots, W_{1N}]}_{\text{Riadok1}}, \underbrace{[W_{21}, \dots, W_{2N}]}_{\text{Riadok2}}, \dots, \underbrace{[W_{N1}, \dots, W_{N,N-1}]}_{\text{RiadokN}} \quad (4)$$

FCM s  $N$  plne prepojenými konceptmi (tj. každý jeden koncept interakcie je spojený so všetkými ďalšími konceptmi) zodpovedá  $N(N - 1)$ -dimenzionálnemu problému. Ak niektoré prepojenie chýba, potom ich zodpovedajúce váhy sú rovné nule a možno ich vynechať, čím sa znižuje rozmernosť problému. Takýto prípad je najčastejší, pretože FCM poskytované odborníkmi sú len zriedka plne pripojené.

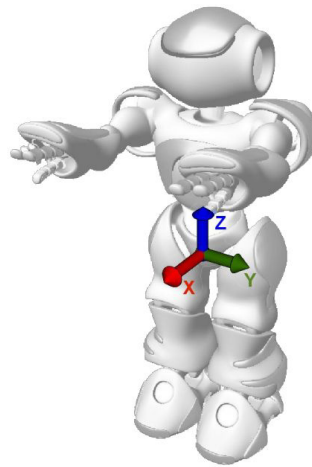
### B. Úprava rozhrania B-Human

Úprava spočíva v generovaní riadiacich bodov pre Bézierovú krivku pomocou fuzzy kognitívnej mapy.

Napríklad ako vstup kognitívnej mapy bude vzdialenosť brány a na výstupe budú riadiace body, ktoré budú určovať silu kopu vzhľadom na danú vzdialenosť. Kognitívna mapa musí byť navrhnutá tak, aby generovala hladký priebeh krivky. Daný výpočet sa musí uskutočniť ešte pred samotným kopom.

Pred samotným kopom sa modul správania nachádza v stave *kick*, kde sa rozhoduje, či daný kop bude vykonaný alebo nie. Ak áno spustí sa funkcia *kickForward*, do ktorej vstupujú parametre polohy lopty. Úprava bude spočívať v tom, že ako parametre budú vstupovať aj poloha brány a polohy prekážok, ktoré pomocou FKM vygenerujú body riadiacej krivky pre kop.

Body riadiacej krivky sú reprezentované trojrozmerným vektorom. Hodnota *x*-ovej súradnice udáva veľkosť smeru nohy dopredu alebo dozadu vzhľadom na robota, hodnota *y*-ovej súradnice udáva veľkosť smeru nohy doľava alebo doprava vzhľadom na robota, hodnota *z*-ovej súradnice udáva veľkosť smeru nohy nahor alebo nadol vzhľadom na robota, tak ako je zobrazené na obr. 3.



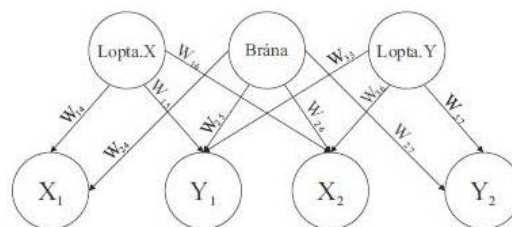
Obr. 3 Vizualizácia systému relatívnych súradníc robota.

Intervaly pre jednotlivé hodnoty súradníc sú v tab. 1.

Tabuľka 1  
 Intervaly pre jednotlivé hodnoty súradníc

	X	Y	Z
MIN	-100	-150	-245
MAX	100	150	70

FKM pre generovanie riadiacich bodov môže vyzerať nasledovne (obr. 4):



Obr. 4 Príklad fuzzy kognitívnej mapy pre generovanie riadiacich bodov.

Na vstupe kognitívnej mapy sa nachádza poloha lopty a vzdialenosť od brány. Koncept *Lopta.X* predstavuje hodnotu *x*-ovej súradnice lopty relatívne k robotovi a koncept *Lopta.Y* predstavuje hodnotu *y*-ovej súradnice lopty relatívne k robotovi. Koncept *Brána* predstavuje relatívnu vzdialenosť brány od robota.

Na výstupe sa nachádzajú jednotlivé parametre krivky. Koncepty  $X_1$  a  $Y_1$  zodpovedajú krivke vo fáze tri a Koncepty  $X_2$  a  $Y_2$  zodpovedajú krivke vo fáze štyri.



Fáza tri zodpovedá ako vysoko zodvihne robot nohu a posunie ju za seba na hodnoty  $X_1$  a  $Y_1$ . Fáza štyri zodpovedá ako vysoko zodvihne robot nohu a posunie ju pred seba na hodnoty  $X_2$  a  $Y_2$ . Rozdiel hodnôt v jednotlivých fázach zodpovedá sile kopu.

#### IV. VYHODNOTENIE POKUSOV

Pokusy boli vykonané v simulátore SimRobot a na reálnom robotovi Nao. Na začiatku experimentu boli experimentálne nájdené optimálne hodnoty konceptov v simulátore SimRobot. Následne prebehla adaptácia fuzzy kognitívnej mapy na dané hodnoty konceptov. Po úspešnej adaptácii bola FKM otestovaná v simulátore a následne aj na reálnom robotovi.

Na vstupe kognitívnej mapy sa nachádza poloha lopty a vzdialenosť od brány. Koncept *Lopta.X* predstavuje hodnotu  $x$ -ovej súradnice lopty relatívne k robotovi a koncept *Lopta.Y* predstavuje hodnotu  $y$ -ovej súradnice lopty relatívne k robotovi. Koncept Brána predstavuje relatívnu vzdialenosť brány od robota.

Na výstupe sa nachádzajú jednotlivé parametre krivky. Koncepty  $X_1$  a  $Y_1$  zodpovedajú krivke vo fáze tri a Koncepty  $X_2$  a  $Y_2$  zodpovedajú krivke vo fáze štyri.

Fáza tri zodpovedá ako vysoko zodvihne robot nohu a posunie ju za seba na hodnoty  $X_1$  a  $Y_1$ . Fáza štyri zodpovedá ako vysoko zodvihne robot nohu a posunie ju pred seba na hodnoty  $X_2$  a  $Y_2$ . Rozdiel hodnôt v jednotlivých fázach zodpovedá sile kopu.

Výsledná FKM má štruktúru zhodnú s obr. 4 a jej váhy nadobúdajú hodnoty podľa (5).

$$\begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.24 & -0.59 & -0.91 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.25 & -0.38 & 0.00 & -0.84 \\ 0.00 & 0.00 & 0.00 & 0.18 & 0.45 & 0.38 & 0.64 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.23 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.58 \\ 0.00 & 0.00 & 0.00 & 0.38 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.34 & 0.00 & 0.00 \end{pmatrix}. \quad (5)$$

Hodnoty konceptov:

Jednotlivé intervaly sú transformované na interval  $[0.0; 1.0]$

Tabuľka 2  
 Hodnoty ustálených konceptov v simulátore

<i>Lopta.X</i>	Brána	<i>Lopta.Y</i>	$X_1$	$Y_1$	$X_2$	$Y_2$	$\Delta X$	$\Delta Y$
0.50	0.50	0.50	0.30	0.50	0.55	0.55	0.20	0.05
0.40	0.60	0.10	0.35	0.50	0.45	0.55	0.15	0.05
0.10	0.10	1.00	0.00	0.10	0.35	0.20	0.35	0.10
0.20	0.30	0.60	0.10	0.30	0.30	0.35	0.20	0.05
1.00	1.00	0.50	0.10	0.90	0.90	0.95	0.80	0.05

Hodnoty konceptov z reálneho robota

Tabuľka 3  
 Hodnoty ustálených konceptov na reálnom robotovi

<i>Lopta.X</i>	Brána	<i>Lopta.Y</i>	$X_1$	$Y_1$	$X_2$	$Y_2$	$\Delta X$	$\Delta Y$
0.48	0.55	0.48	0.25	0.48	0.56	0.54	0.31	0.06
0.38	0.62	0.02	0.30	0.57	0.43	0.49	0.13	-0.08
0.04	0.07	0.94	0.04	0.16	0.32	0.23	0.28	0.07
0.22	0.33	0.63	0.08	0.28	0.28	0.31	0.20	0.03
1.00	0.98	0.47	0.04	0.87	0.84	0.90	0.80	0.03

Z daných tabuliek vyplýva, že robot silno kopne do lopty, preto že rozdiel  $\Delta X$  prekročí hranicu 0.20.

#### V. ZÁVER

Pokusy potvrdili, že použitie fuzzy kognitívnych máp funguje len pri jednoduchých problémoch, ktoré sa dajú linearizovať.

Fuzzy kognitívna mapa generovala optimálne riadiace body pre inverznú kinematiku. Ako vstupný parameter vstupovala poloha lopty a vzdialenosť brány. Vzdialenosť brány ovplyvňovala všetky výstupné koncepty. Keď bol robot pri bráne jeho kop bol slabší ako keď bol od nej ďalej.

Dana metóda fuzzy kognitívnych máp sa ukázala ako vhodná pre opis a simuláciu jednoduchých problémov. Táto práca dokazuje, že je možné generovať dynamické kopy pomocou umelej inteligencie v reálnom čase.

V budúcnosti by bolo vhodné zaoberať sa prihrávkou medzi robotmi.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] T. Laue, T. Röfer, K. Gillmann, F. Wenk, C. Graf, and T. Kastner, “B-human 2011 - eliminating game delays,” in *RoboCup*, ser. Lecture Notes in Computer Science, T. Röfer, N. M. Mayer, J. Savage, and U. Saranlı, Eds., vol. 7416. Springer, 2011, pp. 25–36.
- [2] J. Müller, T. Laue, and T. Röfer, “Kicking a ball - modeling complex dynamic motions for humanoid robots,” in *RobuCup*, ser. Lecture Notes in Computer Science, J. R. del Solar, E. Chown, and P.-G. Plöger, Eds., vol. 6556. Springer, 2010, pp. 109–120.
- [3] B. Kosko, “Fuzzy cognitive maps,” *International Journal of Man-Machine Studies*, vol. 24, no. 1, 1986, pp. 65–75.
- [4] M. Hagiwara, “Extended fuzzy cognitive maps,” in *Fuzzy Systems, 1992., IEEE International Conference on*, 1992, pp. 795–801.
- [5] R. Taber, “Knowledge processing with fuzzy cognitive maps,” *Expert Systems with Applications*, vol. 2, no. 1, 1991, pp. 83 – 87.
- [6] M. Parenthoën, P. Reignier, and J. Tisseau, “Put fuzzy cognitive maps to work in virtual worlds,” in *Proc. The 10th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Melbourne, Australia*, vol. 1, 2001, pp. 252–255.
- [7] E. I. Papageorgiou and J. L. Salmeron, “A review of fuzzy cognitive map research at the last decade,” *IEEE Transactions on Fuzzy Systems*, to be published.
- [8] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena, *Genetic Fuzzy Systems — Evolutionary Tuning and Learning of Fuzzy Knowledge Bases, series Advances in Fuzzy Systems — Applications and Theory*. World Scientific, 2001.

# Simulation of the Vehicle Navigation Using Fuzzy Cognitive Maps

<sup>1</sup>Peter AMRICH, <sup>2</sup>Ján VAŠČÁK

<sup>1,2</sup> Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>peter.amrich@student.tuke.sk, <sup>2</sup>jan.vascak@tuke.sk

**Abstract** — This paper targets on vehicle navigation in unknown environment with obstacles. The navigation is carried out through fuzzy cognitive maps. The paper provides theoretical basis of fuzzy cognitive maps and their use for navigation purposes. Our main goal is to develop navigation simulator. After its completion the program will allow user to choose created fuzzy cognitive map and then visualize behavior of the vehicle according to the selected map. Simulator environment will serve as a tool for experiments realization. The aim of the experiments is to find functional links between changes of parameters of the map and navigation of the vehicle. To evaluate whether fuzzy cognitive maps are suitable for this task, first the manual navigation was realized. Afterwards the results were compared.

**Keywords** — Fuzzy cognitive maps, navigation, simulation, visualization tools

## I. INTRODUCTION

Vehicle navigation is a problem, which is still solved by people. Automatic systems, capable of accomplish this goal, are requested because of number of reasons. It is, for example, conservation of human resources, more effective fuel consumption, or improving transportation planning. Therefore, new methods and approaches to solving the navigation problem are required for further research in the field.

There are a variety of methods for solving navigation problems. The goal of this paper is to describe and implement approach using fuzzy cognitive maps (FCM) for navigation purposes. We believe that this contribution will be beneficial to the reader and to the research in the navigation field.

## II. THEORY OF FUZZY COGNITIVE MAPS

### A. Fuzzy cognitive maps

FCM is an oriented graph. Its nodes represent notions, and the connections between them represent relations. The theory of FCM is well-known. The definition of FCM is more general and extensive [1]. Following brief explanation is a description of FCM used solely in our thesis.

Notions can be set to interval  $\langle 0, 1 \rangle$ . Where 0 means that notion is not valid, 1 means it is valid. The values in between characterize the uncertainty. The relations can be set from the interval  $\langle -1, 1 \rangle$ . They represent knowledge of an expert in the field.

For purpose of our experiments realization following FCM was used. There were several input nodes or notions. Input values from sensors were fuzzified first using membership functions of the nodes. Then they were set to the corresponding nodes of the FCM yet in the required interval. Than new step of FCM state was calculated. Eventually output nodes or notions of the FCM were defuzzified again using membership functions of the corresponding nodes. And the obtained output values were used in navigation for steering the vehicle and determining whether it should stop.

Advantage of using FCM is simple description of the relations in complex dynamic system [1]. Example of such system is navigation of the vehicle between obstacles.

### B. Examples of fuzzy cognitive maps use

FCMs can be used in solving number of problems. As mentioned they are intended to use for description of complex dynamic systems. Except the navigation, they can be and were successfully used in other fields, such as medicine. In [2] is described example of FCM use in

diagnosing bladder cancer. The success rate of diagnostic system was 80% – 90%.

Application of FCM is also suitable for purposes of regulation [3]. In robotics FCM can be used to stabilize movement of a robot or to implement passing football between two robots. The advantage of FCM is that they are capable of working with uncertain information (e.g. from sensors).

In [4] there is described navigation utilizing FCM. The goal was to use a robot to find a cup in unknown environment with obstacles. Parameters of the cup were specified in advance. The Situation Assessment Framework (SAF) was used too. Its job was to generate subgoals which were assessed by FCM. Then the robot decided what it should do.

The similar task to ours is described in article [5]. Artificial Potential Field (APF) method was used to simulate navigation in 2D environment. The obtained results were used to test FCM based navigation. A FCM was created first by defining its nodes and connections. Then parameters of the particular FCM were iteratively changed. To accomplish this, genetic algorithms (GA) were used. After specific number of iterations the result was that FCM copied the behavior of APF method to extent up to 90%.

### III. IMPLEMENTATION

In our thesis we used already available FCM-Editor to create and edit FCM used in navigation. The program which we created was a 3D visualization with implemented physics to simulate real world. It allows user to create environment with obstacles and set target and vehicle position, manually navigate the vehicle to the target and then start the FCM navigation and compare trajectories of both.

All obstacles are represented as walls with constant height of 3 meters and width of 0.2 meters (in metric system of physics simulation). No other obstacles are expected. The vehicle is 3 meters long and 1.5 meters wide. The maximum radius of steering is  $\pm 30^\circ$ . The vehicle is of back-wheel-drive type.

The important issue is also sensors implementation. Every sensor represents angular sector of 0.4 radians and range of 8 meters. There are 6 of them directed forward (Fig. 1). The sensors should provide the closest distance to any obstacle in range and in specified angular sector. Our implementation is carried out using rays. Every sensor has 4 rays. The disadvantage may be that sometimes an obstacle can be in a so called blind angle of the sensor and therefore it doesn't have to be noticed.

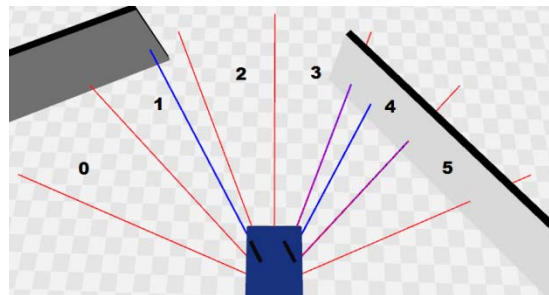


Fig. 1 Sensors of the vehicle. Red lines are borders of the sensors. Blue lines are visualization of the closest distance to obstacles.

### IV. EXPERIMENTS

We used a single thoroughly designed and set FCM for testing of the navigation behavior (Fig. 2, Fig. 3). The FCM was then tested on a variety of scenes with different obstacles and different starting positions of the vehicle. Thereby we were able to prove versatility of using FCMs in navigation of the vehicle.

Nodes of the map represent the notions. Input nodes are b0 – b3 and b8 – b17. Values of these nodes will be set by physics simulation. Then next step of the FCM will be calculated and so output nodes of steering b4 – b6 and node b7 determining whether the vehicle should stop will be set. These will be used to control the vehicle in the simulation.

The goal is to gain the same or similar results of the FCM behavior to the behavior of the human controlled navigation. Therefore prior to the FCM navigation the manual navigation is realized. Total count of 9 experiments will be carried out. Each experiment will have its scene with obstacles. In the scene 3 tests with different starting positions will be made. The example of the output of comparing the results represents Fig. 4.

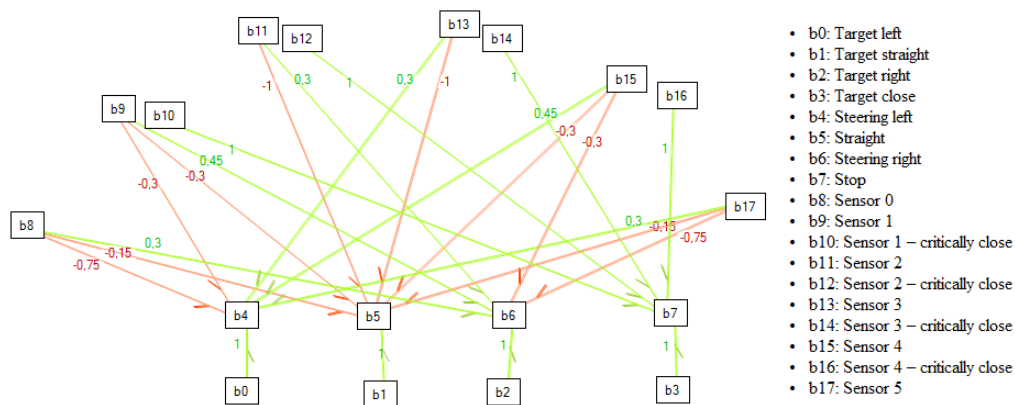


Fig. 2 FCM used in navigation.

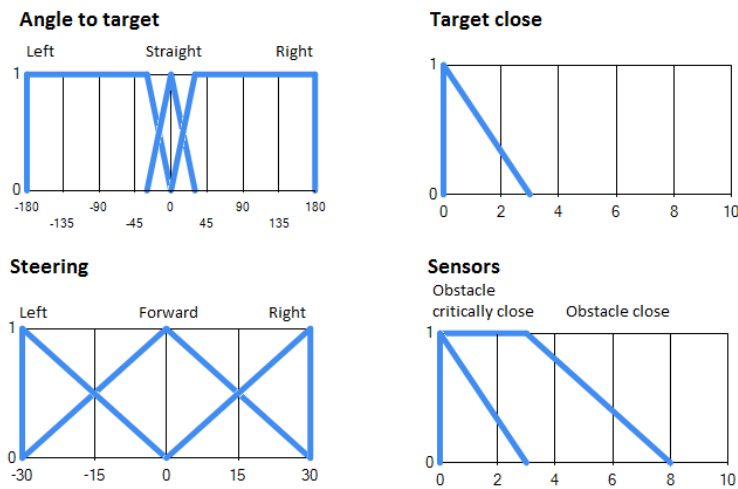


Fig. 3 Membership functions of the nodes in the used FCM.

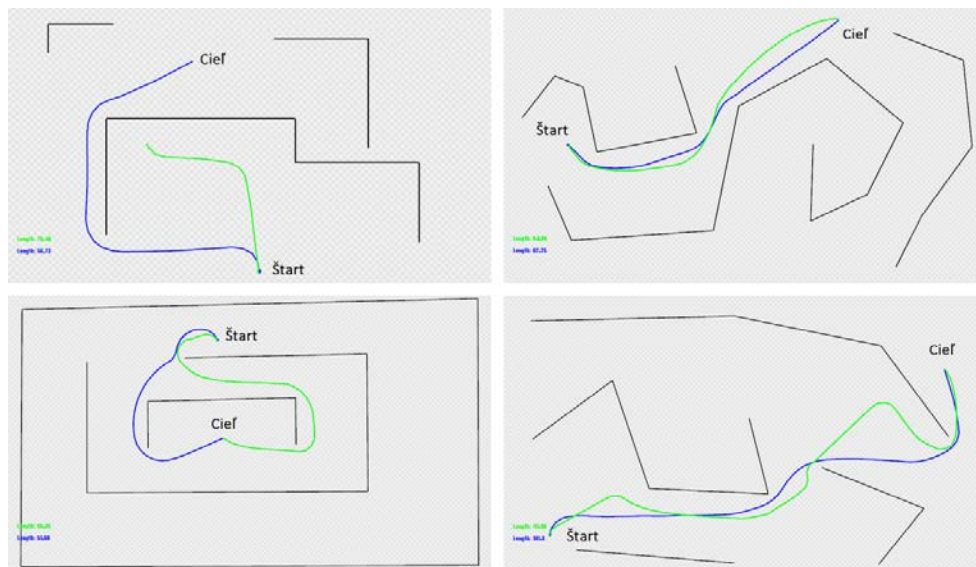


Fig. 4 Examples of experiments outputs. Green color represents FCM navigation, blue color manual navigation.

## V. CONCLUSION

The results of all evaluated experiments are shown in Table 1. The last column represents comparison of the manual and FCM navigations. Symbol 'CN' means that FCM navigation was not able to find the target. This is the result of the implementation of the navigation, where vehicle cannot reverse. Therefore it can come into so called robotic trap, where all sensors report obstacle ahead of them and the vehicle is unable to decide where to go.

Table 1  
Results of experiments

Experiment	Test number	Trajectory length		Comparison
		Manual	FCM	
1	1	44.6	48.4	-3.8
	2	67.6	72.9	-5.3
	3	28.1	27.3	0.8
2	1	58.7	CN	CN
	2	46.6	CN	CN
	3	31.2	31.0	0.2
3	1	88.5	CN	CN
	2	62.3	65.0	-2.7
	3	57.1	56.6	0.5
4	1	75.5	77.8	-2.3
	2	66.6	71.8	-5.2
	3	96.4	100.5	-4.1
5	1	101.1	165.5	-64.4
	2	68.8	70.4	-1.6
	3	51.7	66.4	-14.7
6	1	101.3	112.0	-10.7
	2	76.8	77.8	-1.0
	3	109.9	110.0	-0.1
7	1	73.4	97.4	-24.0
	2	74.7	74.6	0.1
	3	115.6	114.6	1.0
8	1	46.4	46.2	0.2
	2	34.8	45.6	-10.8
	3	36.1	37.9	-1.8
9	1	39.6	CN	CN
	2	29.3	27.3	2.0
	3	28.7	27.5	1.2

In the most cases the FCM navigation was slightly worse or slightly better than the manual navigation. The exceptions were when the path was not found, which could be prevented by implementing reverse movement of the vehicle. And several very long FCM navigation paths (e.g. experiment 5, test 1) which appeared when the vehicle chose wrong turn on the crossroads. We believe that these values are results of short range of the vehicle sensors and also can be prevented. The important factor is also the FCM parameters setting, which we believe also can be improved by implementing similar technique mentioned in the last example in chapter II.B.

#### ACKNOWLEDGMENT

This work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF.

#### REFERENCES

- [1] J. Vaščák, "Approaches in Adaptation of Fuzzy Cognitive Maps for Navigation Purposes," in *8th IEEE International Symposium on Applied Machine Intelligence and Informatics (SAMII)*, 2010, pp. 31-36..
- [2] P. P. Groumpos, "Fuzzy cognitive maps: basic theories and their applications in medical problems," in *19th Mediterr. Conf. Control Autom.*, 2011, pp. 1490-1497.
- [3] R. Michna, *Adaptácia fuzzy kognitívnych máp pomocou evolučných algoritmov pre potreby robotického futbalu*, 2013.
- [4] N. Ganganath, M. Walker, and H. Leung, "Fuzzy Cognitive Map Based Situation Assessment Framework for Navigation Goal Detection," in *2013 IEEE Int. Conf. Syst. Man, Cybern.*, 2013, pp. 1444-1449.
- [5] O. Motlagh, S. H. Tang, A. R. Ramli, N. Ismail, and D. N. Nia, "A Novel Approach to Motion Modeling using Fuzzy Cognitive Map and Artificial Potential Fields," in *11th International Conference on Control Automation Robotics & Vision (ICARCV)*, 2010, pp. 7-10.

# Adaptácia fuzzy kognitívnych máp pomocou algoritmu PSO v robotickom futbale

<sup>1</sup>Patrik Sabol, <sup>2</sup>Ján Vaščák

<sup>1,2</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská Republika

<sup>1</sup>patrik.sabol@student.tuke.sk, <sup>2</sup>jan.vascak@tuke.sk

**Abstrakt** — V tejto práci je predstavené využitie algoritmu Particle Swarm Optimization (PSO) pri úlohách v robotickom futbale. Predstavený je koncept prihrávok medzi robotmi, v ktorom sú využité fuzzy kognitívne mapy, ktoré boli adaptované pomocou algoritmu PSO. Experimentovalo sa s dvoma variantmi PSO, pričom sa testovali rôzne parametre. Najlepšie ohodnotená fuzzy kognitívna mapa bola použitá v systéme prihrávok. Prihrávky boli testované pomocou simulácii humanoidného robota NAO a na reálnych robotoch NAO pomocou rozhrania RoboCup tímu B-Human.

**Kľúčové slová** – fuzzy kognitívna mapa, optimalizácia rojom častíc, prihrávky, inteligencia roja, robotický futbal

## I. ÚVOD

Futbal je kolektívna hra, kde úspešnosť dosiahnutia gólu závisí od vhodnej pozície hráča voči súperovej bránke. To isté platí aj pri RoboCup zápase SPL ligy humanoidných robotov. Pri týchto zápasoch často dochádza k situáciám, že hráč s loptou má malú šancu dosiahnuť gól. Riešením tejto situácie by mohla byť prihrávka adresovaná na robota s väčšou šancou na skórovanie. Avšak aby sa predišlo zbytočnej strate lopty, dôležitá je úspešnosť prihrávky, ktorá závisí od sily a smeru prihrávky.

Princíp prihrávky spočíva v správnom kope do lopty. V [1] navrhli systém kopy, ktorý využíva fuzzy kognitívne mapy (FKM) adaptované pomocou algoritmu Particle Swarm Optimization (PSO). Vstupom FKM bola vzdialenosť brány a poloha lopty relatívne k robotovi. Výstupom FKM boli riadiace body pre krivky, ktoré opisujú jednotlivé fázy pohybu robota.

Cieľom práce bolo upraviť FKM pre problém prihrávky a adaptovať ju pomocou rôznych modifikácií algoritmu PSO. Následne najlepšie ohodnotená FKM sa použila v systéme prihrávok a otestovala na simulátore a na humanoidnom robotovi NAO.

## II. PRINCÍPY ALGORITMU PARTICLE SWARM OPTIMIZATION

### A. Princípy algoritmu PSO

Základom algoritmu PSO je roj, ktorý sa skladá z istého počtu častíc. Tento roj sa pohybuje v  $n$ -rozmernom priestore prehľadávania, pričom každá častica reprezentuje jedno potenciálne riešenie, ktoré je ohodnotené fitness funkciou. Každá častica má pamäť, v ktorej si ukladá svoju najlepšiu pozíciu v rámci priestoru prehľadávania, *pbest* (*personal best*), a pozíciu najlepšej častice z celého roja, *gbest* (*global best*). Tento princíp, zdieľanie skúseností medzi časticami, je základ úspechu algoritmus PSO. Táto vlastnosť sa využíva pri výpočte rýchlosti, ktorou sa častica pohybuje v priestore. Týmto sa roj postupne sústreďuje do oblastí s najlepšimi riešeniami [2].

Postupom času bolo vyvinutých niekoľko verzií tohto algoritmu. Prvotná verzia je opísaná nasledujúcimi vzťahmi:

$$v_{id}(t+1) = v_{id}(t) + c_1 R_1 (p_{id}(t) - x_{id}(t)) + c_2 R_2 (p_{gd}(t) - x_{id}(t)), \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1), \quad (2)$$

kde:

- $v_{id}(t)$  predstavuje rýchlosť  $i$ -tej častice v  $d$ -tej dimenzii priestoru prehľadávania v predchádzajúcej iterácii,
- $x_{id}(t)$  predstavuje pozíciu  $i$ -tej častice v  $d$ -tej dimenzii priestoru prehľadávania v predchádzajúcej iterácii,
- $p_{id}$  predstavuje historicky najlepšiu pozíciu  $i$ -tej častice v  $d$ -tej dimenzii,
- $p_{gd}$  predstavuje pozíciu najlepšej častice v rámci celého roja v  $d$ -tej dimenzii,
- $R_1$  a  $R_2$  predstavujú dva  $n$ -rozmerné vektory s náhodnými číslami z intervalu  $\langle 0.0, 1.0 \rangle$ , ktoré zavádzajú užitočnú náhodnosť do prehľadávacej stratégie,
- $c_1$  a  $c_2$  sú pozitívne konštantné váhové parametre, taktiež nazývané *kognitívne* a *sociálne* parametre, ktoré riadia pohyb každej častice na základe jej individuálnej najlepšej pozície alebo najlepšej pozície roja [3].

Jedna z modifikácií tohto algoritmu vznikla pridaním *zotrvačnej váhy* ( $\omega$ ), ktorá bola zavedená preto, aby v začiatkových fázach optimalizácie bolo uprednostňované globálne prehľadanie, aby sa prehľadal čo najväčší priestor prehľadávania a aby sa ku koncu optimalizačného procesu preferovalo lokálne prehľadanie pre detailnejšie preskúmanie oblastí, ktoré boli nájdené v skorších fázach optimalizácie. Preto bol vzorec (1) upravený na nasledovný tvar:

$$v_{id}(t+1) = \omega(t) * v_{id}(t) + c_1 R_1 (p_{id}(t) - x_{id}(t)) + c_2 R_2 (p_{gd}(t) - x_{id}(t)), \quad (3)$$

kde *zotrvačná váha* je vyjadrená ako

$$\omega(t) = \omega_{hi} - (\omega_{hi} - \omega_{low}) \frac{t}{Tmax}, \quad (4)$$

pričom  $t$  je počítadlo iterácie,  $Tmax$  je maximálny počet iterácií, po ktorých algoritmus skončí,  $\omega_{hi}$  predstavuje inicializačnú hodnotu  $\omega$ , a  $\omega_{low}$  predstavuje minimálnu hodnotu  $\omega$ , ktorú nadobudne, keď  $t = Tmax$  [3].

Ďalšia z modifikácií vznikla pridaním *faktora zúženia* ( $\chi$ ). Tento model je v súčasnosti považovaný za štandard algoritmu PSO a jeho tvar je definovaný nasledujúcou rovnicou:

$$v_{id}(t+1) = \chi [v_{id}(t) + c_1 R_1 (p_{id}(t) - x_{id}(t)) + c_2 R_2 (p_{gd}(t) - x_{id}(t))], \quad (5)$$

kde *faktor zúženia* ( $\chi$ ) je vyjadrený ako

$$\chi = \frac{2}{\left| 2 - \Phi - \sqrt{\Phi^2 - 4\Phi} \right|}, \quad (6)$$

pre  $\Phi > 4$ , kde  $\Phi = c_1 + c_2$  [3].

### B. Využitie algoritmu PSO v úlohách robotického futbalu

Algoritmus PSO pre svoju nízku výpočtovú a časovú náročnosť našiel využitie aj pri úlohách robotického futbalu, kde je potrebná rýchlosť a stabilita robotov.

V [4] využili PSO pri optimalizácii parametrov chôdze humanoídného robota Kondo KHR-1. Optimalizovala sa chôdza do piatich rôznych smerov, pričom ako fitness funkcia sa brala prejdená vzdialenosť za čas. Najlepší optimalizačný progres sa ukázal pri chôdzi vpred, pričom najvyššia dosiahnutá rýchlosť bola 17 cm/sek.

V [5] použili algoritmus PSO pri optimalizácii parametrov metódy filtrácie častíc pre potreby samolokalizácie robota na futbalovom ihrisku. Úlohou bolo nájsť také parametre, pri ktorých lokalizačná chyba bola čo najmenšia, teda fitness funkcia bola definovaná ako hodnota lokalizačnej chyby. Ako tréningové dáta poslúžil 10-minútový záznam zápasu. Po optimalizácii parametrov celková presnosť samolokalizácie robota stúpala. Tieto hodnoty parametrov sa neskôr použili pri zápasoch RoboCup SPL Lígy.



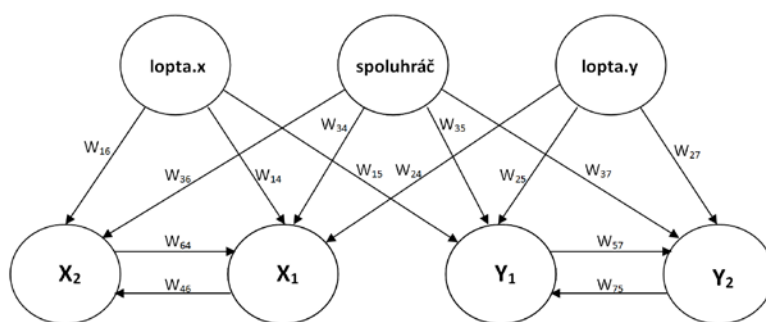
### III. NAVRHOVANÝ PRÍSTUP

#### A. Návrh fuzzy kognitívnej mapy pre prihrávku

Pri návrhu FKM bolo potrebné určiť vstupné a výstupné koncepty tak, aby presne popisovali riešený problém. Vstupné koncepty predstavovali informácie, ktoré môže robot získať zo svojho okolia, konkrétne vzdialenosť spoluhráča a súradnice lopty relatívne k robotovi. Výstupné koncepty predstavujú parametre, ktoré popisujú kop robota do lopty, presnejšie  $X_1$ ,  $Y_1$ ,  $X_2$ ,  $Y_2$ .

Kop do lopty sa skladá z niekoľkých fáz, kde každá fáza je popísaná Bézierovou krivkou. Tá je reprezentovaná ako trojrozmerný vektor, kde hodnota x-ovej súradnice udáva veľkosť smeru nohy dopredu alebo dozadu vzhľadom na robota, hodnota y-ovej súradnice udáva veľkosť smeru nohy doľava alebo doprava vzhľadom na robota, hodnota z-ovej súradnice udáva veľkosť smeru nohy nahor alebo nadol vzhľadom na robota [1].

Teda výstupné koncepty  $X_1$  a  $Y_1$  zodpovedajú krivke vo fáze tri, ktorá opisuje, ako vysoko robot zdvihne nohu a posunie ju za seba na hodnoty  $X_1$  a  $Y_1$ . Výstupné koncepty  $X_2$  a  $Y_2$  zodpovedajú krivke vo fáze štyri, ktorá opisuje, ako vysoko robot zdvihne nohu a posunie ju pred seba na hodnoty  $X_2$  a  $Y_2$ . Navrhovaná FKM je zobrazená na Obr. 1.



Obr. 1 Použitá fuzzy kognitívna mapa pre problém prihrávky.

#### B. Implementácia prihrávky

Pre implementáciu prihrávky boli využité zdrojové kódy RoboCup tímu B-Human, presnejšie bol využitý modul správania. Pre potreby prihrávky boli vytvorené dva typy správania, *Striker* a *Receiver*. Tieto typy správania sa skladajú z hierarchie stavov.

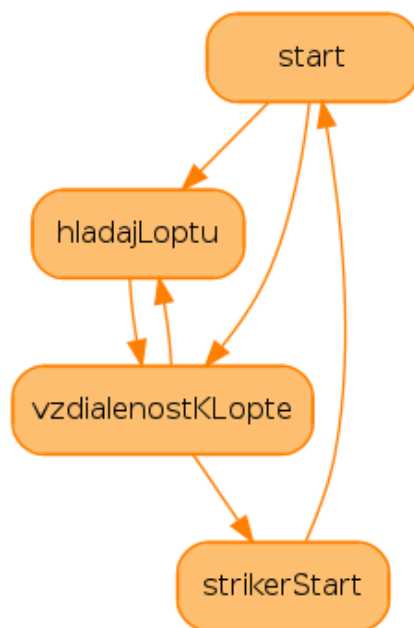
Na začiatku je nastavené správanie *Receiver*, v ktorom sú definované tieto stavy:

- **hladajLoptu** – v prípade, ak robot nemá v dohľade loptu, začne sa otáčať okolo vlastnej osi, kým ju opäť nezbadá.
- **vzdialenosťKlopte** – tu robot vypočíta svoju vzdialenosť od lopty pomocou Euklidovskej vzdialenosti, takisto vypočíta vzdialenosť spoluhráča od lopty, kde údaje ho spoluhráčovi získa cez tímovú komunikáciu. Ak je bližšie k lopte, tak ju prevezme, ak nie, tak ostáva stáť a čaká na prihrávku.
- **strikerStart** – v prípade, ak je robot bližšie ku lopte ako spoluhráč, tak spustí podvoľbu *Striker()*.

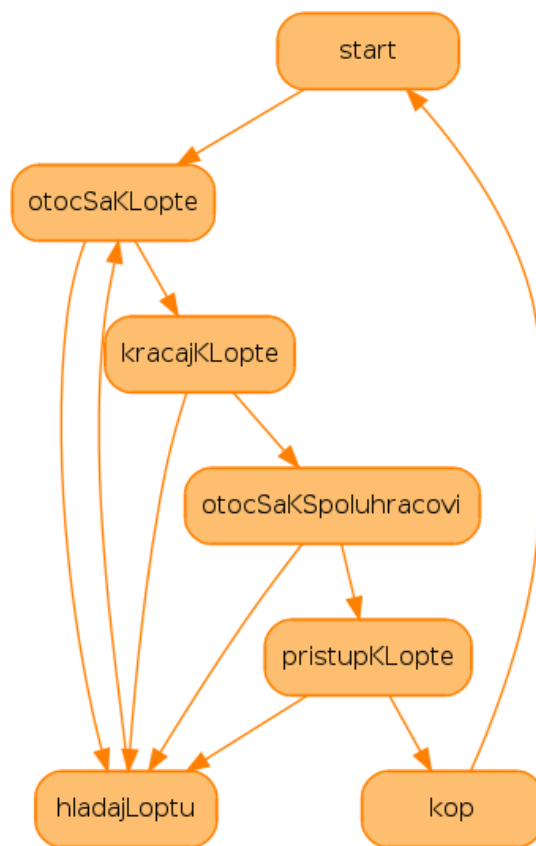
V správaní *Striker* sú definované tieto stavy:

- **otocSaKlopte** – v tomto stave sa vypočíta uhol, na ktorý sa ma robot natočiť, aby bol otočený priamo k lopte.
- **kracajKlopte** – v tomto stave sa vykoná príkaz na chôdzu k lopte
- **otocSaKSpoluhracovi** – tu sa vypočíta uhol, na ktorý sa ma robot natočiť, aby smeroval k spoluhráčovi. Na výpočet uhla sa použili súradnice spoluhráča.
- **prístupKlopte** – v tomto stave robot pristúpi k lopte podľa vypočítaných uhlov s predchádzajúcich stavov a pripraví sa na kop.
- **kop** – v tomto stave prebieha samotný kop do lopty smerom na spoluhráča. Tu sa využíva naučená FKM, ktorej vstupom je poloha lopty a vzdialenosť spoluhráča, ktorá sa vypočíta podľa vzorca (7):

$$\text{vzdialenosťSpoluhrača} = \sqrt{(\text{polohaX} - \text{spoluhračX})^2 + (\text{polohaY} - \text{spoluhračY})^2}. (7)$$



Obr. 2 Grafické znázornenie hierarchie stavov v správaní *Receiver*.



Obr. 3 Grafické znázornenie hierarchie stavov v správaní *Striker*.

#### IV. VYHODNOTENIE EXPERIMENTOV

Experimenty prebiehali v dvoch fázach:

1. Fáza - nájdenie vhodnej FKM pomocou algoritmu PSO
2. Fáza - využitie naučenej FKM pri prihrávke medzi robotmi

##### A. Adaptovanie fuzzy kognitívnej mapy pomocou algoritmu PSO

Pri adaptovaní FKM sa testovali a porovnávali dve varianty algoritmu PSO a *to zotrvačná váha a faktor zúženia*. Experimenty spočívali v menení parametrov týchto variant a sledoval sa počet iterácií, v ktorých dôjde algoritmus do lokálneho, resp. globálneho minima a následne sa porovnávala hodnota fitness funkcie v tomto extréme.

Tabuľka 1  
 Prehľad výsledkov experimentov adaptácie FKM pomocou dvoch variant algoritmu PSO s rôznymi parametrami

	Zotrvačná váha	Fitness	Iterácia
1.	$\omega_{\text{high}} = 1.1 \ \omega_{\text{low}} = 0.1$	3.52093	79
2.	$\omega_{\text{high}} = 1.6 \ \omega_{\text{low}} = 0.4$	3.52093	72
3.	$\omega_{\text{high}} = 2.2 \ \omega_{\text{low}} = 0.8$	3.50821	49
	Faktor zúženia		
1.	$C_1 = C_2 = 2.02 \ \chi = 0.819$	3.50812	74
2.	$C_1 = C_2 = 2.05 \ \chi = 0.729$	3.52903	73
3.	$C_1 = C_2 = 2.6 \ \chi = 0.351$	8.33274	87

Z tabuľky 1 vyplýva, že zvyšovaním parametrov pre *zotrvačnú fázu* sa kvalita nájdeného riešenia výrazne nemenila, ale čas potrebný na nájdenie daného riešenia sa výrazne znižoval.

Zvyšovaním parametrov pre *faktor zúženia* malo opačný efekt, teda kvalita riešenia aj čas potrebný na nájdenie daného riešenia sa postupne zhoršoval.

### B. Vzájomné prihrávanie medzi robotmi

Z predchádzajúcich experimentov sa vybrala najlepšie ohodnotená FKM a použila sa v systéme prihrávok. Prihrávky prebiehali v simulátore SimRobot a na reálnych robotoch NAO.

Experimenty spočívali v striedaní rôznych vzdialeností medzi robotmi, pričom sa sledovala sila a smer prihrávky.

Tabuľka 2  
 Štatistický prehľad vykonaných experimentov v simulátore

Experiment	Vzdialenosť	Počet úspešných prihrávok	Počet uskutočnených prihrávok	Úspešnosť
1.	1 m	4	6	66 %
2.	2,5 m	5	6	83 %
3.	4 m	6	6	100 %
4.	6 m	3	6	50 %

Tabuľka 3  
 Štatistický prehľad vykonaných experimentov na reálnych robotoch

Experiment	Vzdialenosť	Počet úspešných prihrávok	Počet uskutočnených prihrávok	Úspešnosť
1.	1 m	3	6	50 %
2.	2 m	4	6	66 %
3.	3,5 m	0	6	0 %

Z vykonaných experimentov môžeme vidieť (viď tabuľky 2 a 3), že najvyššia úspešnosť prihrávok bola na stredné vzdialenosti. Nižšia úspešnosť prihrávok na krátku vzdialenosť bola spôsobená nedostatočne silným kopom, teda lopta nedošla ku spoluhráčovi, čo malo za dôsledok to, že hráči sa postupne k sebe približovali. Prihrávky na väčšiu vzdialenosť boli úspešné iba v simulátore. Pod neúspešnosťou prihrávok na reálnych robotoch sa podpísal povrch ihriska, ktorý nekládol dostatočný odpor, čoho dôsledkom bolo, že pri silnom kope sa robot otočil okolo vlastnej osi na stojacej nohe, teda loptu vôbec netrafil. Neúspešnosť niektorých prihrávok bola spôsobená aj nedostatočne kalibrovanou kamerou robota, čo spôsobilo nepresný výpočet vzdialeností.

Pri porovnaní výsledkov zo simulátora a z reálu boli lepšie výsledky dosiahnuté v simulátore, pretože simulátor poskytuje ideálne podmienky prostredia, teda ideálny povrch s ideálnym odporom podložky, či presný výpočet vzdialeností.

## V. ZÁVER

Z experimentov vyplynulo, že mierne lepšia variácia algoritmu PSO bola *zotrvačná váha*, keďže dokázala nájsť výsledok v skoršom čase. Tieto experimenty nám ukázali, že algoritmus PSO je veľmi citlivý na hodnotu svojich parametrov.

Táto práca ukázala ďalšiu možnosť využitia algoritmu PSO pri úlohách v robotickom futbale, a to adaptovaním fuzzy kognitívnej mapy pre prihrávku medzi robotmi. Úspešnosť

prihrávk bola pomerne menšia, ale bolo by možné ju zvýšiť použitím presnejšej trénovacej množiny v procese učenia FKM, alebo lepším kalibrovaním kamier robotov.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] R. Michna, "Adaptácia fuzzy kognitívnych máp pomocou evolučných algoritmov pre potreby robotického futbalu," 2013.
- [2] R. C. Eberhart, "Particle Swarm Optimization : Developments , Applications and Resources," pp. 81–86, 1995.
- [3] H. Ahmed, J. Glasgow, and C. Kln, "Swarm Intelligence Concepts , Models and Applications Technical Report 2012-585," no. February, 2012.
- [4] C. Niehaus, T. Röfer, and T. Laue, "Gait optimization on a humanoid robot using particle swarm optimization," in *Humanoid Soccer Robot*, 2007.
- [5] A. Burchardt, T. Laue, and R. Thomas, "Optimizing Particle Filter Parameters for Self-localization," in *Humanoid Soccer Robot* , 2011, pp. 145–156.

# Využitie fuzzy sviečkových grafov pri vyhodnocovaní ekonomických dát

<sup>1</sup>Iveta Krnáčová, <sup>2</sup>Ing. Ján Vaščák

<sup>1,2</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>iveta.krnacova@student.tuke.sk, <sup>2</sup>jan.vascak@tuke.sk

**Abstrakt** — Táto práca sa zaoberá využitím fuzzy logiky vo sviečkových grafoch a ich vyhodnotením pomocou ekonomických dát. Sviečkové grafy sú súčasťou technickej analýzy a sú obľúbené, pretože napomáhajú predpovedať ďalší vývoj trhu. Predpoveď pomocou sviečkových grafov spočíva hlavne v identifikovaní vzorov, ktoré sú zložené z jednej alebo viacerých sviečok. Sviečkové vzory sa skladajú z jednoducho popísaných pravidiel. Zložitejšie na nich je len ich pochopenie, čo nám ktorý vzor predpovedá. Fuzzy logika napomáha sviečkovým vzorom tým, že klasickú logiku rozširuje. Fuzzifikácia sviečkových vzorov v tejto práci je navrhnutá a implementovaná pomocou lingvistických premenných s dodržaním pravidla fuzzy partície. Následne sa vzory vyhodnotia a premietnu do grafov, podľa ktorých ich môžeme vyhodnotiť. V závere práce sú popísané experimenty v ktorých budú porovnávané výsledky fuzzifikovaných sviečkových a klasických vzorov.

**Kľúčové slová** — technická analýza, fuzzy logika, fuzzifikácia, fuzzy sviečkové vzory

## I. ÚVOD

Technická analýza je súčasťou práce každého obchodníka, ktorý v dnešnej dobe obchoduje na akciovom trhu v dnešnej dobe. Medzi základné analýzy, ktoré obchodníci využívajú patrí vyhodnocovanie sviečkových grafov. Sviečkové grafy vznikli pred niekoľkými storočiami v Japonsku. Teória je založená na tom, že trend ktorý nastane môže byť predvídaný pomocou sviečkových vzorov. Signály sviečkových vzorov sú charakteristické pre predvídanie rôznych situácií na trhoch.

Jedným z hlavných cieľov tejto práce je reálnejší popis sviečkových grafov, ich fuzzifikáciu a následné porovnanie klasických a fuzzifikovaných vzorov. Rozdiel medzi klasickou a fuzzy logikou je, že fuzzy logika pracuje na princípe teórie množín snažiac sa zachytiť neistotu a neurčitosť. Tieto systémy môžu, ale nemusia byť riešené aj cez dvojhodnotovú logiku. Nepresnosť, ktorú chceme dosiahnuť u vzorov vyjadříme pomocou lingvistických premenných. Výsledkom experimentov bude vyhodnotenie a porovnanie výsledkov vzorov fuzzifikovaných a nefuzzifikovaných.

## II. TECHNICKÁ ANALÝZA A SVIEČKOVÝ GRAF

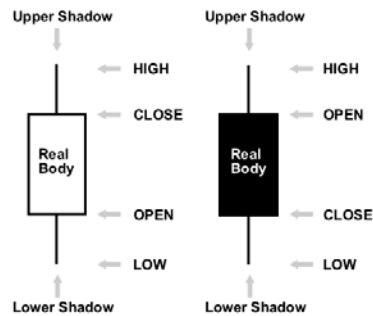
### A. Úvod do technickej analýzy

Technická analýza pomáha určiť kedy je ten správny čas obchodovať. Jej neodmysliteľnou súčasťou sú grafy a využívajú sa hlavne na zistenie, kedy je ten správny moment predat' alebo nakúpiť akciu či komoditu. Počíta s tým, že ceny sa pohybujú iracionálne a ich pohyb je ovplyvnený „vonkajšími silami“. Využitie technickej analýzy je veľmi široké. Môžeme ju prakticky využiť na všetky finančné trhy, či už ide o anglické akcie alebo komodity, francúzske akcie atď. [1]. Základné typy grafov, ktoré sa používajú pri technickej analýze na burze sú: Čiarový graf, Úsečkový graf, Sviečkový graf a Point and Figure graf [2].

### B. Sviečkový graf

So Sviečkovým grafom sa západný svet oboznámil pomerne nedávno, ale v Japonsku bola táto technika známa už dlhšiu dobu. Ide o graf zobrazujúci pohyb ceny v určitom období (najčastejšie jeden deň).

Využíva cenu otváraciu, zatváraciu, najvyššiu a najnižšiu v danom dni. Sviečkový vzor pozostáva z tela, ktorého dĺžka predstavuje rozdiel medzi otváracou a zatváracou cenou a tenkých zvislých čiar nazývaných tiene. Tie sú umiestnené nad a pod telom. Tiene predstavujú najvyššie a najnižšie hodnoty dosiahnuté počas dňa. Tzv. *Bullish* deň s uzatváracou cenou vyššou ako otváracou je znázornený bielym alebo zeleným (dutým) telom, zatiaľ čo tzv. *Bearish* deň s uzatváracou cenou nižšou ako otváracou je znázornený čiernym alebo červeným (vyplneným) telom. Telo sa stáva horizontálnou čiarou, keď sú otváracie a zatváracie ceny rovnaké [3].



Obr. 1 Popis sviečkových vzorov [4].

### C. Používanie sviečkových grafov

Pri posudzovaní pravdepodobnosti obrátenia sa skúsení technickí obchodníci spoliehajú na vzory dvoch alebo viacerých po sebe idúcich sviečok viac ako na jednotlivé sviečky. Trpezlivosť pri technickej analýze je nutná, pretože niekedy sa vzor môže objaviť ako výsledok veľmi krátkeho trendu alebo horizontálneho pohybu, no nemusí byť nikdy potvrdený. V takýchto prípadoch nepotvrdenia je lepšie počkať, aby sa zabránilo strate obchodu až do doby, kým nevznikne nový vzor [3].

Čítanie vzorov alebo sviečkových formácií zo sviečkových grafov je základom obchodovania na burze, keď sa riadime podľa sviečkového typu grafu. Sviečkové vzory sa vyskytujú v rôznych kombináciách. Na základe signálov sa obchodník rozhoduje, či jeho obchodnú pozíciu na trhu otvorí alebo zatvorí. Nerozhoduje sa iba podľa signálov zo sviečkových vzorov, ale používa pri tom aj iné metódy, čiže tieto metódy kombinuje. Sviečkové vzory sa delia na štyri základné skupiny [5]:

1. Bullish vzory
2. Bearish vzory
3. Reverzné vzory
4. Neutrálne vzory

Medzi Bullish vzory v tejto práci boli použité vzory: Morning Star, Piercing Line, Three White Soldiers, Engulfing, Hammer, Harami, Upside Tasuki Gap, Marubozu

Bearish vzory: Evening Star, Tree Black Crows, Engulfing, Hanging Man, Harami, Downside Tasuki Gap, Marubozu.

Z neutrálnych vzorov bol použitý Doji.

### D. Fuzzy logika

Fuzzy logika je podmnožinou konvenčnej (booleovskej) logiky. Uvedený typ logiky je rozšírený o čiastkové pravdivostné hodnoty. Čiastkové údaje sú zobrazené medzi „úplne pravdivými“ a „úplne nepravdivými“ hodnotami. Už samotný názov hovorí, že je to logika, ktorá je iba orientačná, nemusí byť presná. Vyplýva zo skutočnosti, pretože spôsob ľudského uvažovania a prirodzený úsudok majú takisto orientačný charakter [6].

Fuzziifikácia sviečkových vzorov sa vykonáva s využitím takých pojmov, ako sú fuzzy množiny, fuzzy partície a lingvistické premenné.

### III. FUZZIFIKÁCIA SVIEČKOVÝCH VZOROV

#### A. Sviečkové vzory

Sviečkové vzory nie sú úplne presne matematicky zadefinované a môžeme tvrdiť, že sú opísané presne tak, ako má daný konkrétny vzor vyzeráť. Pozostávajú z poznatkov nadobudnutých obchodníkmi. Pre objasnenie problematiky bude uvedený príklad, ako zo slovne zadefinovaných vzorov extrahujeme podmienky pre dané vzory. Pre názornú ukážku bol vybraný vzor Bearish Engulfing.

Slovným opisom je charakteristický Bearish Engulfing tým, že má čierne telo, ktoré je predchádzané menším, bielym telom. Obe telá sa objavia v priebehu rastúceho trendu. Čierna sviečka nemusí nutne pohltiť tieň bielej sviečky, ale pohltí telo samotné.

Pravidlo pre Bearish Engulfing vyzerá takto:

prvý je biely &  
 druhý je čierny &  
 zatváracia cena 1 < otváracia cena 2 &  
 otváracia cena 1 > zatváracia cena 2 &  
 veľkosť prvého je Malá &  
 veľkosť druhého je Veľká &  
 trend je Bearish

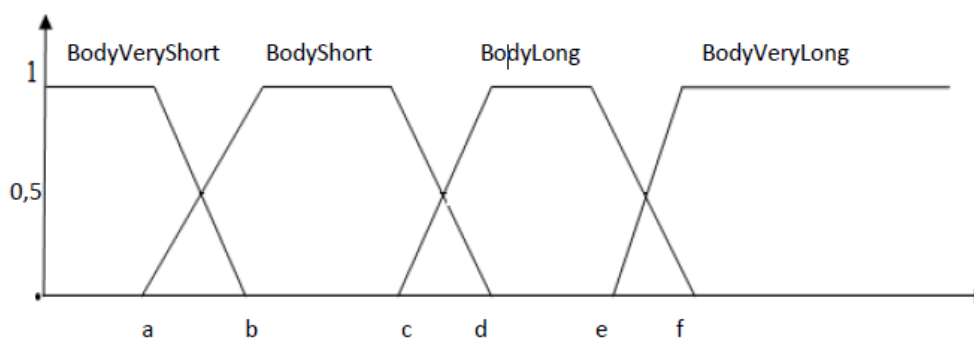
Takáto formácia vie nájsť len ostré hodnoty daných vzorov.

#### B. Fuzzifikácia

Fuzzifikácia sviečkových vzorov, čiže časti, ktoré je možné fuzzifikovať a na ktorých sa fuzzifikácia aj vykonala, sú veľkosti a tieň sviečok. Farbu sviečok sme nefuzzifikovali, lebo je vždy vyjadrená bielou alebo čiernou farbou a ostane dvojrozmerná.

Fuzzifikované sviečkové vzory sú reprezentované fuzzy lingvistickými premennými. Lingvistické premenné sú vyjadrené ako Body Length, Shadow Length, Upper Shadow Length a Lower Shadow Length. V týchto 4 lingvistických premenných sa nachádzajú ďalšie 4 hodnoty, ktoré vyjadrujú jednotlivé fuzzy množiny a určujú veľkosti tela a tieňov. Napr. BodyLength obsahuje fuzzy množiny BodyVeryLong, BodyLong, BodyShort, BodyVeryShort.

Na Obr. 2 (pozri nižšie) je znázornená ukážka funkcie príslušnosti  $\mu(x)$  BodyLength pre výpočet stupňa príslušnosti tela sviečky. Toto znázornenie funkcie príslušnosti je použiteľné aj na lingvistické premenné tieňov sviečok. Dĺžky premenných nie sú pevne stanovené, môže si ich určiť obchodník alebo používateľ sám podľa potreby manuálne tak, aby zodpovedali potrebám cieľov. Pri vytváraní funkcie príslušnosti sa dodržiava pravidlo fuzzy partície.



Obr. 2 Funkcia príslušnosti.

$$\text{Body Very Short } (x) = \begin{cases} 0, & x > b \\ 1, & x \in [\infty; a] \\ \frac{1}{(a-b)} * (x - b), & x \in [a; b] \end{cases} \quad (1)$$

$$\text{Body Short } (x) = \begin{cases} 0, & < a \text{ alebo } > b \\ \frac{1}{(b-a)} * (x - a), & [a; b] \\ 1, & [b, c] \\ \frac{1}{(c-d)} * (x - d), & [c; d] \end{cases} \quad (2)$$

$$\text{Body Long } (x) = \begin{cases} 0, & < c \text{ alebo } > d \\ \frac{1}{(d-c)} * (x - c), & [c; d] \\ 1, & [d, e] \\ \frac{1}{(e-f)} * (x - f), & [e; f] \end{cases} \quad (3)$$

$$\text{Body Very Long } (x) = \begin{cases} 0, & > e \\ \frac{1}{(f-e)} * (x - e), & [f; e] \\ 1, & [f; \infty] \end{cases} \quad (4)$$

#### IV. EXPERIMENTY

Pre experimenty boli vybrané dáta cien zlata a platiny. Ceny sú datované od 14. októbra 2009 do 16. októbra 2013. Experimenty boli vyhodnotené na 16 vzoroch.

Na základe pozorovania sa mohlo začať vyhodnocovať, či daný vzor skutočne nastal alebo predikcia nebola pravdivá. Pri pozorovaní sa muselo rozoznávať, či je daný vzor Bullish alebo Bearish. Autor musel byť oboznámený, či daný vzor patrí do kategórie reverzných vzorov, čiže sa smer chystá zvrátiť alebo je to vzor, ktorý len posilňuje smer Bullish či Bearish. Dôvodom pre udanie vierohodnosti je fakt, že sviečkové vzory sú súčasťou technickej analýzy, ktorá by sa mala kombinovať spolu s minimálne jednou analýzou. Ak je vierohodnosť nízka, môže sa kombinovať dokonca aj s mnohými ďalšími technickými analýzami tak, aby obchodník vedel s väčšou istotou, či obrat nastane alebo nie.

Výsledky vyhodnotenia experimentov boli vložené do Tabuľky 1 (pozri nižšie) a boli vyhodnocované na základe dvoch vzorcov.

Vzorec pre klasickú logiku:

$$OU = \frac{\sum_{i=1}^{n1} UE}{\sum_{i=1}^{n1} UE + \sum_{i=1}^{n2} NE} \quad (5)$$

Vzorec pre fuzzy logiku:

$$FU = \frac{\sum_{i=1}^{n1} SP * UE}{\sum_{i=1}^{n1} SP * UE + \sum_{i=1}^{n2} SP * NE} \quad (6)$$

OU – ostrá úspešnosť

UE – úspešné experimenty

NE – neúspešné experimenty

FU – fuzzy úspešnosť

SP – stupeň príslušnosti

Tabuľka 1  
Vyhodnotenia experimentov.

Názov	Klasická logika- Zlato	Fuzzifikácia č.1- Zlato	Fuzzifikácia č.2- Zlato	Klasická logika- Platina	Fuzzifikácia č.1 - Platina	Fuzzifikácia č.2- Platina
Morning Star	57% z 7	84% z 9	71% z 9	50% z 12	53% z 11	53% z 12
Piercing Line	50% z 2	56% z 4	65% z 4	66% z 3	0% z 0	0% z 0
Three White Soldiers	100% z 1	100% z 2	100% z 1	100% z 3	66% z 3	66% z 3
Engulfing	48% z 42	44% z 41	43% z 41	54% z 54	51% z 50	50% z 57
Hammer	33% z 18	38% z 9	61% z 4	33% z 15	36% z 28	48% z 14
Harami	55% z 36	50% z 34	46% z 37	51% z 39	47% z 36	49% z 42
Upside Tasuki Gap	57% z 7	57% z 7	57% z 7	0% z 2	0% z 2	0% z 2
Marubozu	39% z 56	33% z 156	29% z 89	38% z 39	31% z 107	26 z 57
Evening Star	48% z 31	65% z 32	68% z 35	51% z 45	61% z 37	51% z 44
Tree White	0% z 0	0% z 0	0% z 0	0% z 2	43% z 7	50% z 6
Crows						
Marubozu	25% z 59	31% z 137	28% z 89	35% z 26	34% z 65	28% z 31



Engulfing	44% z 41	50% z 43	49% z 43	41% z 58	47% z 50	51% z 55
Hanging Man	32% z 22	45% z 25	66% z 9	25% z 20	43% z 40	61% z 18
Harami	36% z 22	56% z 31	56% z 35	43% z 29	51% z 34	45% z 40
Downside	0% z 0	0% z 0	0% z 0	50% z 2	50% z 2	50% z 2
Tasiku Gap						
Doji	31% z 91	34% z 252	33% z 201	42% z 151	38% z 253	42% z 201

### A. Výsledky experimentov

Experimenty preukázali, že fuzzifikácia sviečok prebiehala úspešne. Aj keď úspešnosť pravdivosti vzorov nebola až taká vysoká, pohybovala sa okolo 40% - 50%, musíme znova zdôrazniť, že sviečkové vzory sú len súčasťou technickej analýzy.

Na základe výsledkov experimentov môžeme skonštatovať, že vo väčšine sviečkových vzorov bola percentuálna vierohodnosť fuzzifikovaných vzorov lepšia a počet nájdených sviečkových vzorov bol taktiež vyšší. Nastala však aj situácia, že fuzzifikované vzory nedosiahli pri niektorých experimentoch lepšie výsledky ako klasické, či už v úspešnosti vzorov alebo vo vyhľadani celkového počtu vzorov v grafe.

Na základe vykonaných experimentov v práci vyplynulo, že najvierohodnejším zo 16 uvádzaných vzorov je Morning Star. Mal jednu z najlepších percentuálnych vierohodností. Avšak v grafoch sa nevyskytol až tak často ako niektoré iné vzory ako napr. Doji, Marubozu alebo Engulfing.

Vzory Bullish Engulfing a Harami sa stali percentuálne menej vierohodnými po fuzzifikácii ako boli pred fuzzifikáciou. Vo väčšine prípadov sa vyhľadal v grafe aj menší počet vzorov ako pri klasickej logike. Môžeme skonštatovať, že týmto vzorom konkrétna fuzzifikácia nepomohla, ale je určitá pravdepodobnosť, že pri zmene lingvistických parametrov by fuzzifikácia dopadla lepšie.

Dvomi nasledujúcimi vzorom ako sú Upside Tasuki Gap a jeho podobný vzor typu Bearish Downside Tasuki Gap sa ich počty ani vierohodnosť po fuzzifikácii nezmenili. Príčina je, že fuzzifikácia bola implementovaná na veľkosť samotných tiel a ich tieňov. Lenže pri týchto vzoroch nie je v popisoch ani pravidlách určené a dôležité aké veľké majú byť ich telá či tieň, podstatná je ich pozícia. Čiže aj tu bolo potvrdené, že fuzzifikácia prebehla úspešne, keďže počty nájdených vzorov a úspešnosť ostali nezmenené.

Marubozu a Doji sú vzory, kde podstatu vzoru tvorí jedna sviečka a to má za následok, že sa v grafoch vyskytovali častejšie ako ostatné. Počet vzorov sa po fuzzifikácii zvýšil podstatne viac ako pri ostatných skúmaných vzoroch. Pozitívum je, že aj keď boli často v grafe identifikované, ich percentuálna úspešnosť bola na relatívne veľké množstvo nájdených vzorov dosť vysoká.

## V. ZÁVER

V súčasnosti dokážu algoritmy sviečkových vzorov pomerne často predvídať vývoj cien. V dnešnej dobe sú už tieto algoritmy reprezentované pomocou počítačovej techniky. Využitie fuzzy logiky pre účely rozoznávania vzorov pomôže a umožní algoritmom rozpoznať sviečkové vzory podobne ako ich rozoznávajú ľudia. V tejto práci sme rozšírili pravidlá rozoznávania tak, aby veľkosť sviečok, resp. tieňov určoval systém tak, že bude možné určiť, či daná sviečka nie je len malá alebo veľká, ale aby bolo možné povedať, že sviečka je stredne malá alebo stredne veľká. Takýmto spôsobom sa určil stupeň príslušnosti každej veľkosti sviečky a horného alebo dolného tieňa.

Následne sa začali experimenty, v ktorých sme vyhodnocovali konkrétne vzory. Na vyhodnocovanie vzorov boli použité ekonomické dáta z minulosti. Vyhodnocovanie prebiehalo takým spôsobom, že systém našiel daný vzor a na základe pozorovania sa určilo či nájdený vzor predpovedal predikciu správne alebo sa mylil a predikcia nebola správna.

Vykonané experimenty ukázali, že pravidlá upravené fuzzy logikou dosiahli oproti klasickej metóde lepšie výsledky. Len v ojedinelých prípadoch nastali situácie, kedy bola klasickej logika percentuálnejšie úspešnejšia ako tá fuzzifikovaná.

Fuzzy sviečkové vzory podávajú bohatšie informácie o cenách, ktoré daný obchodník skúma. Ďalšou výhodou fuzzy vzorov je, že pri objavení sa sviečkového vzoru na grafe je konkrétnemu vzoru priamo priradená váha, čiže stupeň príslušnosti. Obchodník alebo používateľ takto vie, aká vážnosť prislúcha danému signálu.

Signály, ktoré nám dávajú sviečkové vzory, možno stále považovať len za náznaky, ktoré pomáhajú lepšie sa rozhodovať obchodníkom o tom, či na trhu predat' alebo nakúpiť. Preto pre úspešné obchodovanie obchodníci používajú sviečkové grafy v kombinácii s inými technickými analýzami alebo faktormi.

## POĎAKOVANIE

Výskum podporovaný Národným projektovým grantom pre výskum a vývoj 1/0667/12 „Inkrementálne metódy učenia pre inteligentné systémy“ 2012-2015.

## LITERATÚRA

- [1] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, “KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework,” *Journal of Multiple-Valued Logic and Soft Computing*, vol. 17, no. 2–3, 2011, pp. 255–287.
- [2] K. Aho, “Benchmarking main activation functions in fuzzy cognitive maps,” *Expert Systems Applications*, vol. 36, no. 3, 2009, pp. 5221–5229.
- [3] E. Markar, “Fuzziness in the semantic web: Survey and future directions,” in *Proc. SEKE'08*, 2008, pp. 643–648.
- [4] R. Prado, S. García-Galán, J. Muñoz Expósito, and A. Yuste, “Knowledge acquisition in fuzzy–rule–based systems with particle–swarm optimization,” *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 6, 2010, pp. 1083–1097.
- [5] M. Gavalec and K. Mls, “Trend evaluation in on-line decision making,” in *Proc. KOI 2008*, Pula, Croatia, ser. International Conference on Operational Research, 2008, pp. 267–274.
- [6] J. H. Wilson, *A learning system based on genetic adaptive algorithms*, Ph.D. dissertation, Department of Computer Science, University of Pittsburgh, USA, 1980.

# Energetická Optimalizácia Regulátora Dynamického Systému

<sup>1</sup>Daniela PERDUKOVÁ, <sup>2</sup>Pavol FEDOR

<sup>1,2</sup>Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>daniela.perdukova@tuke.sk, <sup>2</sup>pavol.fedor@tuke.sk

**Abstract** — Predložený príspevok sa zaoberá metodikou návrhu parametrov regulátora s využitím fuzzy logiky, ktorý je pre daný spojitý dynamický systém optimálny z energetického hľadiska pri dodržaní požadovaných dynamických parametrov regulácie. Vhodné parametre regulátora sa hľadajú na základe fuzzy modelu sústavy, získaného jej identifikáciou z nameraných vstupov a výstupov. Postup je verifikovaný na príklade návrhu parametrov PI regulátora pre jednosmerný pohon s nelineárnou záťažou.

**Keywords** — dynamický systém, energetická optimalizácia, fuzzy model, fuzzy regulátor, jednosmerný pohon, nelineárna záťaž

## I. ÚVOD

Jedným zo základných spôsobov, ako zvyšovať energetickú efektívnosť existujúcich alebo novo navrhovaných elektroenergetických zariadení je optimalizácia ich prevádzky nasadzovaním tzv. inteligentných technológií ich riadenia. Návrh regulátorov s pevne zadanými parametrami nie je z tohto hľadiska vždy optimálny, pretože pri ňom nemusia byť presne známe parametre regulovanej sústavy alebo sa tieto môžu meniť počas prevádzky zariadenia (napr. odporu vinutí elektromotorov môžu meniť svoje hodnoty po zahriatí až o 30 %, záťaž pohonu môže byť nelineárna a pod.) [1, 2]. Preto je vhodné identifikovať aktuálne parametre regulovanej sústavy a na ich základe voliť aj aktuálne energeticky optimálne hodnoty parametrov regulátora.

Predkladaný článok testuje jeden z možných spôsobov takéhoto postupu, založený na využití znalostí z oblasti fuzzy systémov. Je tu stručne uvedený postup zostavenia fuzzy modelu dynamického systému a jeho využitie pri optimalizácii parametrov PI regulátora. Uvedená metodika je overená simuláciami v programe MATLAB na pohone s jednosmerným motorom a tiež sú tu uvedené možné energetické úspory pri regulácii takéhoto pohonu.

## II. ZOSTAVENIE FUZZY MODELU DYNAMICKÉHO SYSTÉMU

Dynamický systém, ktorý chceme riadiť, môžeme popísať v spojitom tvare v stavovom priestore rovnicou

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x}(t), t) * \mathbf{x}(t) + \mathbf{B}(\mathbf{x}(t), t) * \mathbf{u}(t) \quad (1)$$

kde  $\mathbf{A}(\mathbf{x}(t), t)$  je matica systému,  $\mathbf{B}(\mathbf{x}(t), t)$  je vstupná matica,  $\mathbf{x}(t)$  je stavový vektor systému a  $\mathbf{u}(t)$  je vektor vstupov do systému (v ďalšom texte budeme tieto matice a vektory označovať zjednodušene  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{x}$ ). V diskretnom tvare môžeme popísať systém (1) rovnicou

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} + \Delta \mathbf{x}_{k-1} \\ \Delta \mathbf{x}_{k-1} &= \mathbf{A}' \mathbf{x}_{k-1} + \mathbf{B}' \mathbf{u}_{k-1} \end{aligned} \quad (2)$$

pričom uvažujeme dobu vzorkovania veličín systému  $T$  a  $k$  je celé číslo, označujúce krok vzorkovania. To znamená, že stav systému v  $k$ -tom kroku je funkciou (všeobecne neznámou a nelineárnou) stavu a vstupov systému v  $(k-1)$  kroku.

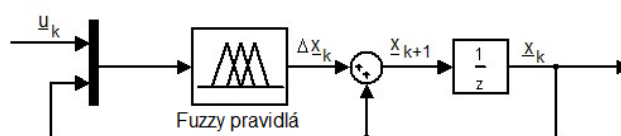
Lahko sa presvedčíme, že ak doba vzorkovania  $T$  spĺňa Shanon-Kotelnikove podmienky, potom približne platí

$$\mathbf{A}' = \mathbf{A} * \mathbf{T}, \mathbf{B}' = \mathbf{B} * \mathbf{T} \quad (3)$$

V literatúre sa uvádza množstvo rôznych štruktúr pre fuzzy modely dynamických systémov [3, 4, 5, 6, 7, 8, 9]. Pri zostavovaní fuzzy modelu dynamickej sústavy (1) sme vychádzali z jej popisu v diskretnom tvare podľa rovnice (2), pričom, ako už bolo uvedené, všeobecne platí, že zmena stavu systému (1) v danom okamihu závisí od jeho vstupu a jeho stavu v predchádzajúcom okamihu. Uvedenú skutočnosť využijeme pri zostavení tzv. stavového fuzzy modelu dynamického systému, kde pre diskretný popis vzťahu medzi stavom  $\mathbf{x}$  a vstupom  $\mathbf{u}$  je možné použiť fuzzy pravidlá v tvare:

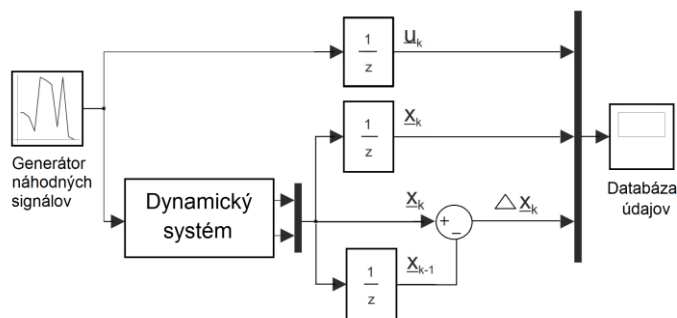
$$\text{IF } u_k \text{ is ... and } x_k \text{ is ... THEN } \Delta y_k \text{ is ..} \quad (4)$$

kde index  $k$  znamená hodnotu veličiny v  $k$ -tom kroku vzorkovania a výrazy  $u_k$  je ...,  $x_k$  je ...,  $\Delta x_k$  je ... predstavujú príslušné fuzzy definičné oblasti danej veličiny. Na základe uvedeného môžeme zostaviť blokovú schému fuzzy modelu dynamickej sústavy (obr. 1), kde blok  $z^{-1}$  predstavuje onekorenie o jeden vzorkovací krok.



Obr. 1 Štruktúra fuzzy modelu diskretného dynamického systému

Pre zostavenie fuzzy modelu potrebujeme nazbierať zo sústavy databázu odpovedajúcich vstupných a výstupných signálov, napr. zapojením podľa obr. 2. Fuzzy model sústavy budeme teda zostavovať na základe nameraných relácií medzi  $[u_k, x_k] \rightarrow \Delta x_k$ , pričom na jej vstup privedieme rôzne hodnoty vstupného signálu tak, aby sme pokryli celý rozsah možných vstupov a tým celý pracovný priestor možných stavov sústavy.



Obr. 2 Zber údajov do databázy pre zostavenie fuzzy modelu

Pre zostavenie uvedeného modelu pre konkrétny dynamický systém, potrebujeme vhodným meraním získať odpovedajúce hodnoty vstupov a výstupov v jednotlivých krokoch vzorkovania napr. vo forme nasledujúcej tabuľky:

Tabuľka 1  
 Forma vzorovej tabuľky databázy údajov pre zostavenie fuzzy modelu

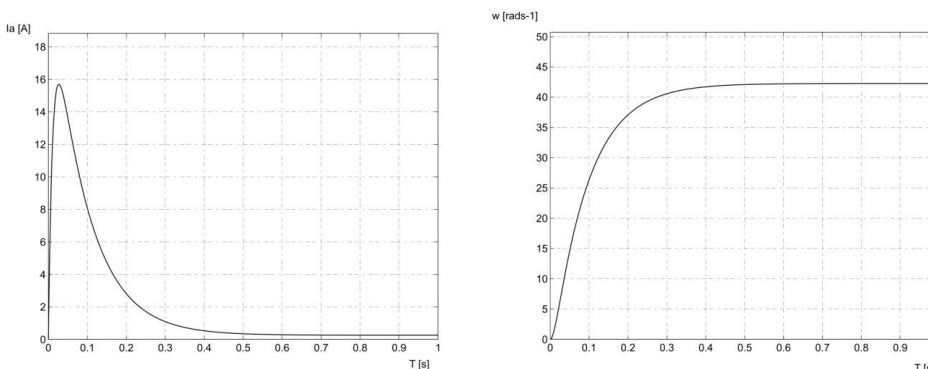
$u_k$	$x_k$	$x_{k+1}$
45	62.9892	0.002
90	62.9894	6.8796
:	:	:
130	123.584	8.9842

Na základe nameranej databázy je možné hľadať FIS štruktúru fuzzy modelu danej sústavy, ktorá najlepšie popisuje namerané relácie. V programovom balíku MATLAB bol na to použitý Fuzzy Toolbox a jeho nástroj Anfisedit.

**Príklad 1:** Ako dynamický systém 2. rádu budeme uvažovať jednosmerný cudzobudený motor s parametrami  $K_A = 0.625 \Omega^{-1}$ ,  $T_A = 0.01$  s,  $c\phi = 0.7$  Vs,  $J = 0,03$  kgm<sup>2</sup>. Ak zvolíme stavové veličiny motora  $x_1 = \omega$  (uhlová rýchlosť motora) a  $x_2 = I_A$  (prúd kotvy motora), potom budú matice jeho stavového popisu v tvare

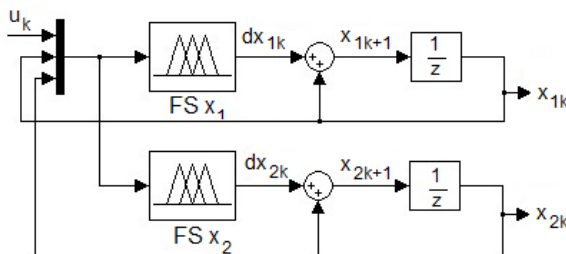
$$\mathbf{A} = \begin{bmatrix} 0 & 22.22 \\ -43.75 & -100 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 \\ 62.5 \end{bmatrix} \quad (5)$$

Pri uvažovaní záťažného momentu ventilátorového typu  $M_Z = k \cdot x_1^2$  ( $k=0.0001$ ) je odozva tohto pohonu na na pripojenie vstupného napätia  $U_A = 30$ V zobrazená na obr. 3.



Obr. 3 Odozva JS motora na skok napätia  $U_A = 30$ V na kotve

Štruktúra fuzzy modelu jednosmerného pohonu, ako dynamického systému 2. rádu, je zobrazená na obr.4. Fuzzy model pozostáva z dvoch fuzzy podsystémov pre jednotlivé stavové veličiny uvažovaného pohonu, pričom každý fuzzy podsystém má tri vstupy a jeden výstup.



Obr. 4 Fuzzy model jednosmerného pohonu

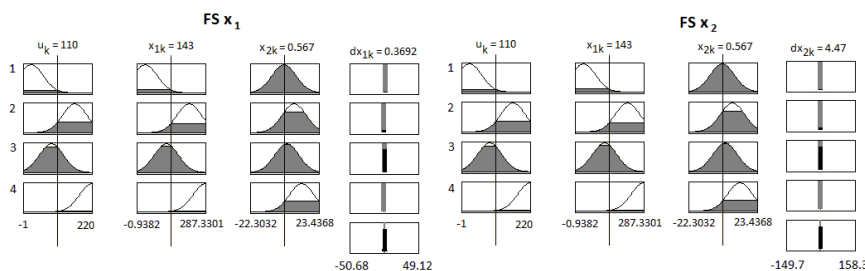
Je zrejmé, že pre vytvorenie fuzzy modelu pre celú pracovnú oblasť pohonu je potrebné namerať hodnoty odpovedajúcich veličín vo všetkých jeho potenciálnych stavoch. V našom prípade sme rozdelili rozsah vstupného napätia od 0 do 220 V na 10 úrovní a do databázy údajov pre zostavenie fuzzy modelu sme namerali požadované hodnoty jednotlivých veličín pre všetky prechody medzi týmito úrovňami (Tab. 2).

Tabuľka 2  
 Štruktúra tabuľky hodnôt potrebných pre zostavenie fuzzy modelu systému 2. rádu

Údaje pre FS $x_1$				Údaje pre FS $x_2$			
$u_k$	$x_{1k}$	$x_{2k}$	$dx_{1k}$	$u_k$	$x_{1k}$	$x_{2k}$	$dx_{2k}$
90.00	123.58	2.18	0.00	90.00	123.58	2.18	0.00
130.00	123.58	2.18	6.10	130.00	123.58	2.18	20.33

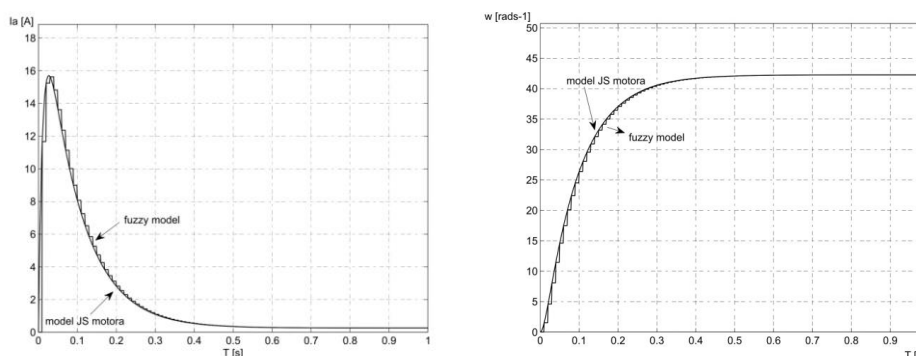
Pomocou nástroja Anfisedit v programe Matlab boli z nameranej databázy údajov (Tab.2) zostavené dva statické fuzzy systémy typu Sugeno, z ktorých každý obsahoval štyri pravidlá. Počet pásem pre fuzzyfikáciu vstupných veličín a tvar funkcií príslušnosti pre obidva fuzzy

systémy sú zrejme z obr. 5.



Obr. 5 Štruktúra fuzzy podsystemov FSx<sub>1</sub> a FSx<sub>2</sub> (fuzzifikácia a pravidlá)

Porovnanie odoziev jednosmerného pohonu s jeho fuzzy modelom na vybraný skok napätia U=50 V je na obr. 6 a ukazuje praktickú zhodu modelu a pohonu v dynamických stavoch.



Obr. 6 Porovnanie dynamiky jednosmerného pohonu a jeho fuzzy modelu

### III. OPTIMALIZÁCIA PARAMETROV REGULÁTORA POHONU

Po vytvorení a overení fuzzy modelu dynamickej sústavy pre jej pracovný priestor je pomocou neho možné hľadať optimálne parametre zvoleného regulátora, najčastejšie typu PI.

Najprv si musíme vymedziť priestor pre rozsah parametrov regulátora, pričom môžeme postupovať aj tak, že v prípade analytickej znalosti systému ich vypočítame klasickým spôsobom, čím získame informáciu o ich počiatočných hodnotách, v okolí ktorých potom definujeme ich rozsah. Pre takto vypočítané parametre regulátora určíme energetickú spotrebu pohybu systému, pričom nás bude zaujímať, či existujú aj iné parametre regulátora, ktoré by boli energeticky optimálne. Za týmto účelom je potrebné cyklicky prechádzať definovanú množinu parametrov regulátora a pre danú konkrétnu dvojicu parametrov a konkrétnu žiadanú hodnotu vstupu vypočítať energetickú spotrebu pohybu sústavy. Z uvedených údajov sa potom vylúčia tie, ktoré nevyhovujú stanoveným hraničným podmienkam a súčasne sa vyberú parametre, pri ktorých je pohyb energeticky najmenej náročný. Uvedený algoritmus môžeme realizovať napríklad prostredníctvom vytvoreného m-file v programe Matlab.

**Príklad 2:** Pre jednosmerný pohon z Príkladu 1 sme definovali nasledujúce hraničné podmienky pre dynamiku jeho pohybu:

- Čas nábehu ±10 % žiadaného času nábehu (0,5 s),
- Otáčky v čase nábehu ±10 % žiadaných otáčok,
- Maximálny prúd pri nábehu nesmie prekročiť dvojnásobok nominálneho prúdu 2I<sub>n</sub>.

Kritériom pre výber optimálnych parametrov PI regulátora bolo minimum spotrebovanej energie podľa vzorca

$$E = \int_0^t U_A I_A dt \tag{6}$$

kde t je čas simulácie.

Najprv sme vypočítali parametre PI regulátora jednosmerného pohonu bez uvažovania záťaže klasickým spôsobom podľa kritéria optimálneho modulu (K<sub>P</sub> = 3.15 a K<sub>I</sub> = 35.01). Uvedené hodnoty nám poslúžili na definovanie rozsahu hodnôt parametrov regulátora (K<sub>P</sub> = 2:0.1:4, K<sub>I</sub> =

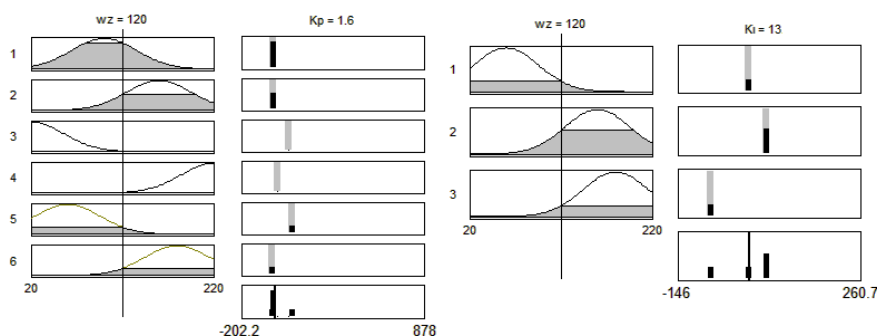
10:5:50). V rámci takto vymedzeného priestoru sme potom hľadali pre konkrétnu hodnotu vstupu takú dvojicu parametrov, aby bolo splnené kritérium pre minimum spotrebovanej energie podľa rovnice (6) a tiež hraničné podmienky pre dynamiku uvažovaného systému.

Výsledky optimalizačného m-file sú zobrazené v Tabuľke 3.

Tabuľka 3  
 Optimálne parametre PI regulátora v rámci jeho pracovného priestoru

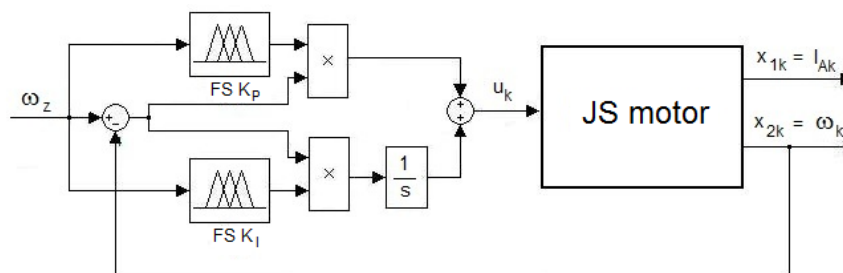
$\omega_z[\text{rads}^{-1}]$	$K_P$	$K_I$	$E_{\text{opt}} [\text{J}]$	$E[\text{J}]$	Úspora
20	1.5	12	29.98	39.09	30.38 %
40	1.5	12	61.83	43.30	42.79 %
60	1.55	12	94.56	72.67	30.12 %
80	1.6	12	147.00	118.34	24.21 %
100	1.6	12	156.03	197.07	26.83 %
120	1.6	13	233.40	288.40	23.56 %
140	1.5	14	351.45	410.50	16.80 %
160	1.5	14	469.08	569.00	21.30 %
180	1.55	13	656.04	770.20	17.40 %
200	1.5	14	884.7	1020.10	15.30 %
220	1.5	14	1134.03	1356.30	19.60 %

Z uvedenej tabuľky vyplýva, že parametre PI regulátora  $K_P$ ,  $K_I$  (vzhľadom na to, že sa jedná o systém s nelineárnou záťažou) je potrebné nastavovať podľa veľkosti žiadanej hodnoty otáčok pohonu. Na základe nameraných relácií medzi  $\omega_z \rightarrow K_P$  a  $\omega_z \rightarrow K_I$  v Tab. 3 vytvoríme pomocou nástroja Anfisedit v programe Matlab dva jednoduché statické fuzzy systémy typu Sugeno (FS  $K_P$  a FS  $K_I$ ), ktorých štruktúra (počet pravidiel, spôsob fuzzyfikácie a tvar funkcií príslušnosti) je zrejماً z obr. 7.



Obr. 7 Štruktúra fuzzy systémov  $K_P$  a  $K_I$ .

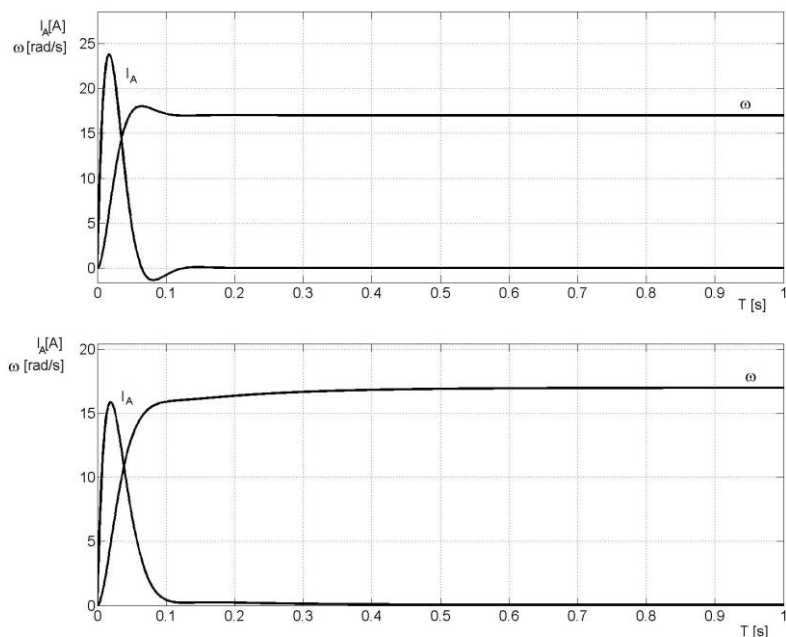
Výsledná bloková schéma energeticky optimálneho regulátora pre jednosmerný pohon s nelineárnou záťažou je na obr. 8, kde fuzzy systémy FS  $K_P$  a FS  $K_I$  zabezpečujú výber parametrov PI regulátora na základe žiadanej hodnoty otáčok tak, aby bol systém optimálny z hľadiska spotrebovanej energie podľa kritéria (6), a súčasne aby vyhovoval požadovaným hraničným kritériám.



Obr. 8 Optimálny fuzzy regulátor jednosmerného pohonu s nelineárnou záťažou

Odozvy štandardného PI regulátora, navrhnutého podľa kritéria optimálneho modulu z parametrov pohonu bez uvažovania záťaže a PI regulátora s optimalizovanými parametrami pre skok napätia na kotve motora  $U_A = 30 \text{ V}$  sú uvedené na obr. 9. Z energetického porovnania

oboch prechodových dejov vyplýva, že v tomto konkrétnom prípade je možné ušetriť cca. 31.27% elektrického príkonu uvažovaného pohonu.



Obr. 9 Odozvy pohonu so štandardným a optimalizovaným PI regulátorom pre skok napätia na kotve motora  $U_A=30V$

#### IV. ZÁVER

V článku je stručne popísaný postup návrhu parametrov regulátora, ktorý zabezpečí energeticky optimálnu dynamiku regulácie spojitého dynamického systému pri definovaných hraničných požiadavkách. Tento postup je verifikovaný na konkrétnom príklade jednosmerného pohonu s nelineárnou záťažou. Výber parametrov regulátora je robený na základe fuzzy modelu sústavy získaného jej identifikáciou z nameraných vstupov a výstupov. V porovnaní so štandardne navrhnutým PI regulátorom s konštantnými parametrami pre celú pracovnú oblasť pohonu je v danom konkrétnom prípade možné ušetriť v priemere až 15,3 % elektrickej energie pri každom dynamickom pohybe pohonu. Táto metodika návrhu regulátorov bude teda vhodná hlavne pre pohony s častými zmenami pracovných bodov, napr. manipulátorov výrobných liniek, pohony elektromobilov a pod. Postup je samozrejme použiteľný aj pre polohovú reguláciu pohonu, prípadne pohony s inými typmi motorov.

#### POĎAKOVANIE

Publikované s podporou projektu KEGA 011TUKE-4/2013.

#### LITERATÚRA

- [1] J. Vittek, S. J. Dodds, *Forced Dynamics Control of Electric Drives*. Žilina: Žilinská univerzita v Žiline, 2003.
- [2] M.M. Krishan, "Fuzzy Sliding Mode Control with MRAC Technique Applied to an Induction Motor Drives", *International Review of Automatic Control (IREACO)*, Praise Worthy Prize, 1(1), 2008, pp.42-48
- [3] R. Babuška, H.B. Verbruggen, H. Hellendoorn, "Promising Fuzzy Modeling and Control Methodologies for Industrial Applications", In: *Proceedings European Symposium on Intelligent Techniques ESIT'99*, Greece, Crete, AB-02, 1999
- [4] T. Takagi, M. Sugeno, "Fuzzy identification of systems and its application to modeling and control", *IEEE Trans. Systems, Man and Cybernetics*, 15, 1985, pp.116-132
- [5] P. Brandstetter, L. Stepanec, "Fuzzy Logic in Vector Controlled Induction Motor Drive", In: *Proc. of EPE 2003 Conference*, Toulouse, France, 2003
- [6] O. Bourebia, K. Belarbi, "Fuzzy Generalized Predictive Control for Nonlinear Systems with Coordination Technique", *International Review of Automatic Control (IREACO)*, Praise Worthy Prize, 2(1), 2008, pp. 169-176
- [7] R. Babuska, *Fuzzy Modeling for Control*. Kluwer Academic Publishers, Boston, 1998
- [8] Y. Lu Xu, "Fuzzy model identification and self-learning for dynamic systems", *IEEE Trans. Systems Man Cybernet.*, SMC-17, 4, 1997
- [9] J. Žilková, J. Timko, S. Kover, "DTC on Induction motor Drive", *Acta Technica CSAV*, 4 (56), 2011, pp. 419-431



# Regulácia servopohonov Dynamixel pri pohybe robotického manipulátora

<sup>1</sup>Marek VACEK, <sup>2</sup>Jaroslava ŽILKOVÁ, <sup>3</sup>Martin LEŠO

<sup>1,2,3</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika

<sup>1</sup>marek.vacek@tuke.sk, <sup>2</sup>jaroslava.zilkova@tuke.sk, <sup>2</sup>martin.leso@tuke.sk

**Abstrakt** — Článok sa zaoberá možnosťami riadenia pohybu päťkĺbového robotického manipulátora, ktorý pozostáva zo servopohonov Dynamixel. Je opísaný postup odosielania užívateľom definovaných údajov, ktoré sa ukladajú do jednotlivých adries riadiacej jednotky servopohonov. Podľa nich vie servopohon meniť parametre pohybu, ako sú rýchlosť, dynamika rozbehu a dobehu, presnosť dosiahnutia určenej polohy a ďalšie. Článok sa taktiež venuje porovnaniu nastavení jednotlivých parametrov a následne ich vzájomnému porovnaniu.

**Kľúčové slová** — dynamixel, SDK, robotický manipulátor, servopohon.

## I. ÚVOD

Dynamixel je vysokovýkonný systém vyvíjaný firmou Robotis, ktorý je určený pre robotické aplikácie a dá sa riadiť prostredníctvom digitálnych paketov. Systém sa skladá zo samotného servopohonu (jednosmerného motora), zabudovanej prevodovky, snímačov rôznych parametrov a riadiacej jednotky, ktorá sa stará o komunikáciu medzi systémom Dynamixel a PC (prostredníctvom komunikačnej jednotky).

K výhodám tohto systému patrí okamžitá spätná väzba o dosiahnutí požadovanej polohy, preťaženie alebo o prehriatí, prípadne možnosť ďalšej komunikácie medzi sebou – servopohony sú zapojené do série. V ovládacom programe je zadané, ktorý parameter je určený pre ktorý servopohon. Signál sa potom vysiela z pohonu do pohonu (smer komunikácie – prvý motor – druhý motor – tretí motor...) pričom požiadavka programu sa vykonáva iba na zvolenom servomotore (definovanie ID motora).

## II. OVLÁDANIE SERVOPOHONOV

Ovládanie servopohonov sa vykonáva prostredníctvom načítavania a zapisovania údajov do ovládacej tabuľky. Tieto údaje sú uložené pod preddefinovanými adresami v pamäti servopohonu. Údaje v pamäti ostávajú zapísané, pokiaľ je servopohon pripojený k napájaciemu napätiu. Spomedzi vyše päťdesiat dostupných nastavení môžeme ako príklad uviesť:

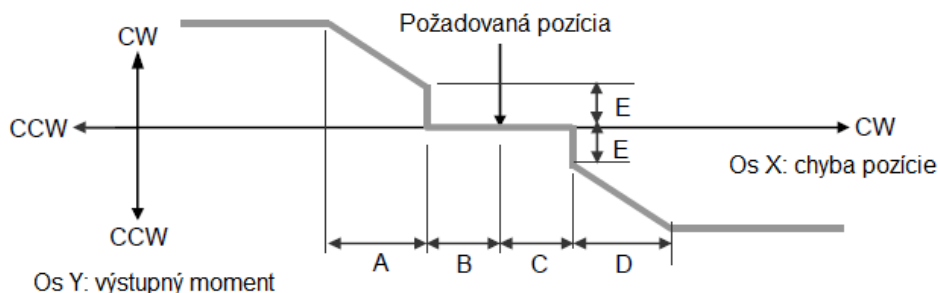
- 0x1E – dolný bajt požadovaného natočenia
- 0x1F – horný bajt požadovaného natočenia
- 0x20 – dolný bajt požadovanej uhlovej rýchlosti
- 0x21 – horný bajt požadovanej uhlovej rýchlosti
- 0x28 – dolný bajt aktuálneho zaťaženia pôsobiaceho na servopohon
- 0x29 – horný bajt aktuálneho zaťaženia pôsobiaceho na servopohon
- 0x2A – aktuálne napätie servopohonu
- 0x2B – aktuálna teplota servopohonu

Všetky typy servopohonov Dynamixel používajú rovnaké konfiguračné adresy. Významnou odlišnosťou medzi použitými sériami servopohonov (AX a MX) v ovládanom robotickom manipulátore je nastavenie ich dynamických vlastností.

Pre sériu servomotorov s označením MX sa nastavenie potrebnej dynamiky definuje PID regulátorom natočenia, jednotlivé adresy pamäte sú:

- 0x1A – derivačná zložka regulátora
- 0x1B – integračná zložka regulátora
- 0x1C – proporcionálna zložka regulátora

Pre sériu servopohonov AX sa nastavujú do RAM pamäte parametre zobrazené na Obr.1.



Obr.1 Nastavenie dynamiky servopohonov typu AX

- A: sklon CCW (adresa 0x1D)
- B: vôľa CCW (adresa 0x1B)
- C: vôľa CW (adresa 0x1A)
- D: sklon CW (adresa 0x1C)
- E: hodnota momentu v kľudovom stave (adresa 0x30, 0x31)

CW – ClockWise (otáčanie v smere hodinových ručičiek)  
 CCW – CounterClockWise (otáčanie proti smeru hodinových ručičiek)

Zmena všetkých parametrov servopohonu prebieha pomocou Dynamixel Software Development Kit (SDK), ktorý okrem zmeny parametrov umožňuje aj riadenie komunikácie medzi servopohonmi Dynamixel a počítačom. Knižnica obsahuje napríklad príkazy na:

- nadviazanie komunikácie
- nastavenie komunikačnej rýchlosti
- definovanie sériového portu
- overenie úspešnej komunikácie
- zápis požadovaných premenných do definovaných adries

### III. ROBOTICKÝ MANIPULÁTOR

Navrhnutý robotický manipulátor sa skladá z piatich servopohonov Dynamixel. Uloženie jednotlivých servopohonov a rozmery jednotlivých častí je možné opísať Denavit-Hartenbergovými parametrami, ktoré sú spolu s typmi pohonov pre jednotlivé kĺby a rozsahy natočenia jednotlivých kĺbov robotického ramena uvedené v Tab.1. Uvedené rozsahy jednotlivých kĺbov zohľadňujú fyzické možnosti natočenia kĺbov robota a rozsahy natočenia jednotlivých servopohonov. Denavit-Hartenbergov parameter  $\theta$  reprezentuje premennú veličinu, ktorá opisuje natočenie kĺbov. Nulová hodnota parametrov  $\theta$  predstavuje počiatočnú pozíciu robota, ku ktorej sa vzťahujú limity natočenia.

Tab. 1 Opis robotického ramena

Kĺb	Pohon	Rozsah	d	a	$\alpha$	$\theta$
1.	MX-28T	-174° až 186°	136,9 mm	0 mm	-90°	0°
2.	MX-64T	-90° až 90°	0 mm	201,8 mm	0°	0°
3.	MX-64T	0° až 180°	0 mm	200 mm	-90°	0°
4.	AX-12A	-150° až 150°	0 mm	18,5 mm	0°	0°
5.	AX-12A	0° až 180°	80,9 mm	0 mm	0°	0°

Pre presný a plynulý chod robotického ramena je nutné nastaviť správne hodnoty dynamických parametrov servopohonov. Pri výbere jednotlivých servopohonov je vhodné zohľadniť nadobúdaciú cenu, konštrukciu a uvážiť životnosť servopohonov pri častom použití.

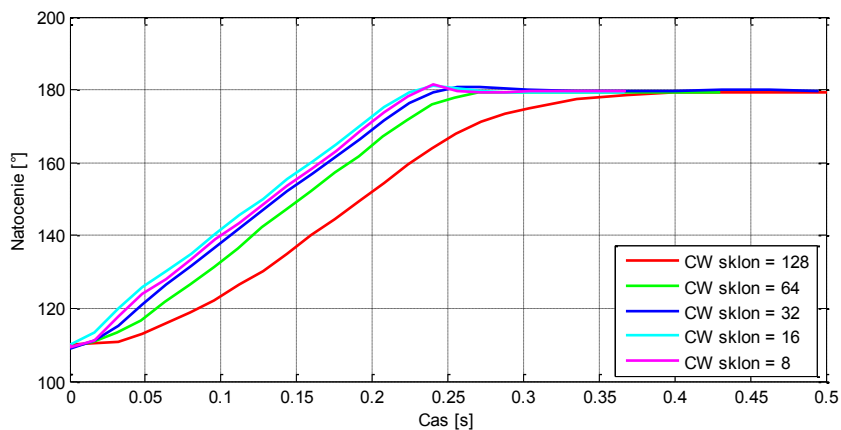
### IV. REGULÁCIA JEDNOTLIVÝCH SERVOPOHONOV

#### A. Nastavenie dynamických parametrov pre pohony série AX

Na Obr.1 je uvedený popis jednotlivých parametrov ovplyvňujúcich dynamiku servopohonov.

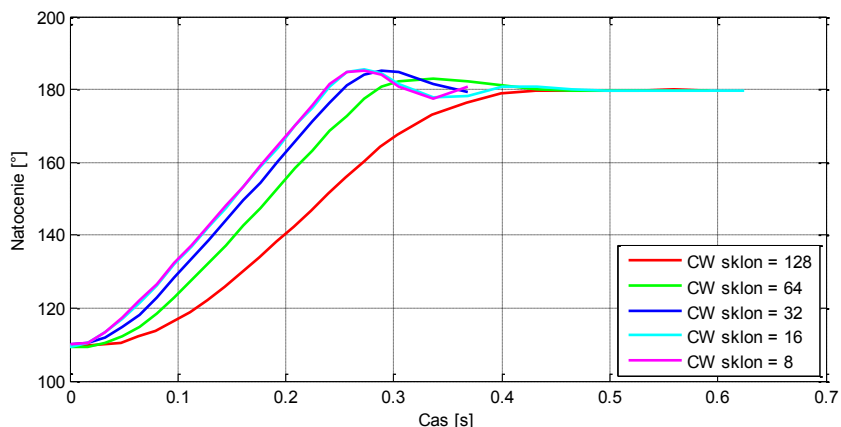
Ako vyplýva z obrázka, veľkosť momentu servopohonu je konštantne závislá od veľkosti odchýlky pozície servopohonu. Vzhľadom na to, že požadovanú reguláciu momentu bolo potrebné aplikovať pre oba smery otáčania rovnako, dvojica parametrov CW sklon a CWW sklon bola nastavená vždy na rovnaké hodnoty. Parametre vôle CW a vôle CWW boli nastavené na minimálnu hodnotu 1, aby sa dosiahlo čo najpresnejšie natočenie servopohonu.

Na Obr.2 sú zobrazené priebehy natočenia servopohonu v čase, pri zmene požadovanej hodnoty natočenia pre rôzne nastavenie parametrov sklonu CW. Dané meranie prebehlo na nezaťaženom servopohone bez prídavného momentu zotrvačnosti, pri nastavení 50% maximálnej uhlovej rýchlosti.



Obr.2 Nastavenie dynamiky servopohonov série AX (bez prídavného momentu zotrvačnosti)

V prípade zväčšenia momentu zotrvačnosti hriadeľa servopohonu, pri nastavení 50% maximálnej uhlovej rýchlosti, je priebeh znázornený na Obr.3

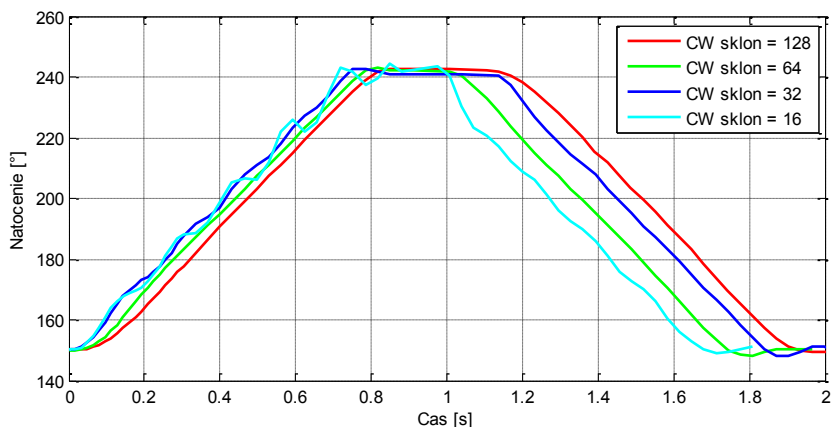


Obr.3 Nastavenie dynamiky servopohonov série AX (zväčšenie momentu zotrvačnosti)

Posledné meranie bolo uskutočnené na robotickom manipulátore pre štvrtý kĺb. Počiatočná pozícia bola nastavená na nulovú polohu, pričom rameno sa malo natočiť o 90°, zotrvať v požadovanej polohe a následne sa vrátiť na počiatočnú pozíciu (Obr.4).

Z meraní vyplýva, že vhodné nastavenie sklonu CW je na hodnotu 128, pretože nedochádza k prekmitom pri dosahovaní požadovanej polohy. Tým bol dosiahnutý plynulý pohyb robotického ramena. Hlavnou nevýhodou takéhoto nastavenia je neschopnosť dosiahnuť presne požadovanú polohu, čo vyplýva z Obr.1.

Vzhľadom k vôľam a životnosti prevodoviek servopohonov je takéto nastavenie najvýhodnejšie.



Obr.4 Nastavenie dynamiky servopohonov série AX (meranie pre štvrtý kĺb robotického manipulátora)

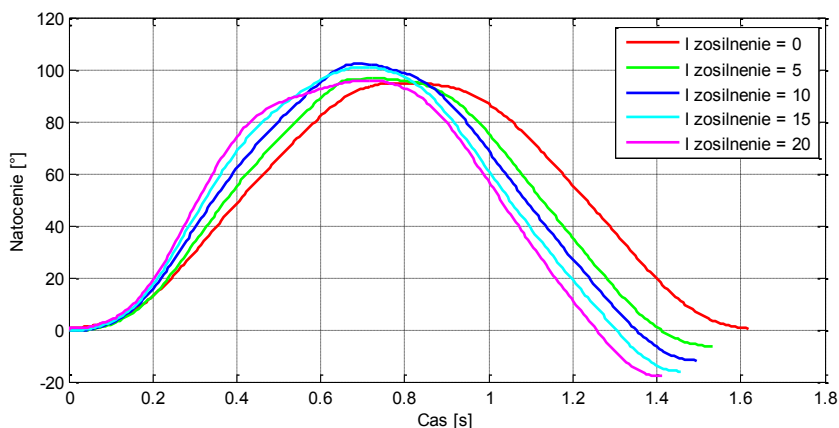
### B. Nastavenie dynamických parametrov pre pohony série MX

Servopohony rady MX používajú na regulovanie požadovanej polohy PID regulátor. Tým je možné odstrániť hlavnú nevýhodu servopohonov rady AX, a to nedosiahnutie požadovanej polohy. PID regulátor sa skladá z proporcionálnej, integračnej a derivačnej časti. Pri našom meraní sa menila hodnota nastavenia iba proporcionálnej a integračnej zložky zosilnenia regulátora. Derivačná zložka zosilnenia bola nastavená na nulu. Veľkosť proporcionálnej zložky je možné určiť z Tab.2, ktorá vyjadruje vzťah medzi nastavenou hodnotou sklonu a nastavením proporcionálnej zložky zosilnenia regulátora.

Tab. 2 Porovnanie nastavenia sklonu ku nastaveniu proporcionálnej zložky PID regulátora

Sklon	P-zosilnenie
8	128
16	64
32	32
64	16
128	8

Nastavením zosilnenia proporcionálnej zložky zosilnenia regulátora na hodnotu 8 dostávame dynamiku rady servopohonov MX podobnú rade AX. Dodatočným nastavením zosilnenia integračnej zložky je možné vykompenzovať vzniknutú odchýlku natočenia servopohonu pri jeho trvalom zaťažení.



Obr.5 Vplyv integračnej zložky regulátora

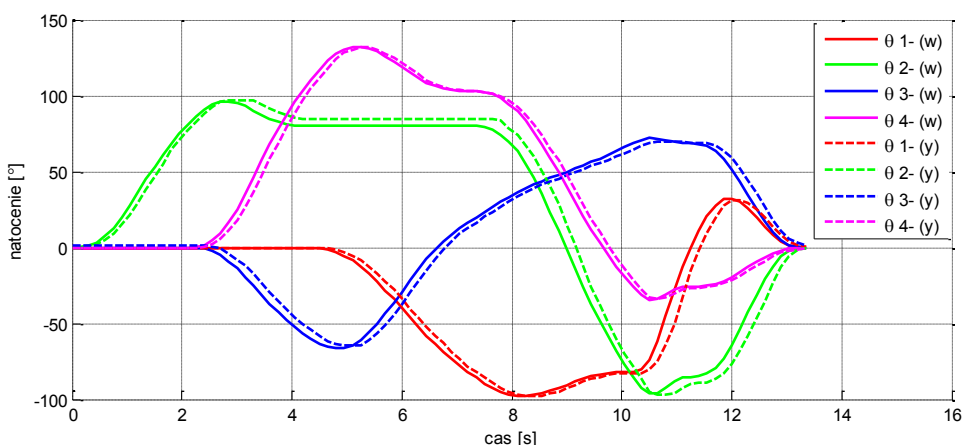
Merania boli vykonané na druhom kĺbe robotického manipulátora obdobne ako pri predchádzajúcom meraní. Počas merania boli namerané hodnoty zosilnenia integračnej zložky regulátora (Obr.5).

Všeobecne sa so zvyšujúcim zosilnením integračnej zložky zvyšuje dynamika pohonu a rýchlosť vyregulovania regulačnej odchýlky. V prípade nastavenia príliš vysokej hodnoty zosilnenia dochádza ku kmitaniu.

## V. ZÁVER

Tento článok sa zaoberá možnosťami riadenia pohybu päťkĺbového robotického ramena pozostávajúceho zo servopohonov Dynamixel, bližšie z rady AX a MX. Všetky riadiace príkazy sa do servopohonov posielajú pomocou preddefinovaných adries, ktoré sú pre všetky série rovnaké. Rozdiely medzi sériami AX a MX sú hlavne v type ich konštrukcie a inej riadiacej jednotke. Kým parametre pohybu servopohonu AX sú definované iba ako hodnota sklonu a vôle CW a CCW, pri sérii MX je k dispozícii plnohodnotný PID regulátor. Pomocou tohto regulátora je možné dosiahnuť presnejšie hodnoty natočenia kĺbu robotického manipulátora do užívateľom definovanej polohy.

Na Obr.6 je znázornené porovnanie požadovanej hodnoty natočenia štyroch servopohonov k nameraným hodnotám (keďže piaty servopohon natáča iba efektor robotického manipulátora, meranie bolo uskutočnené bez neho). Z obrázka je možné pozorovať, že robotický manipulátor mierne zaostáva za požadovanými hodnotami. V prípade, že servopohon je pod väčšou záťažou (hodnota jeho momentu je nad 70%), je schopný dosiahnuť požadované natočenie, ale s malou odchýlkou (približne  $1,5^\circ$ ).



Obr.6 Porovnanie požadovaného natočenia servopohonov (w) ku skutočnému (y)

## POĎAKOVANIE

Publikované s podporou projektu KEGA 011TUKE-4/2013.

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-0185-10.

## REFERENCIE

- [1] Reza N. Jazar: Theory of Applied Robotics. Springer New York Dordrecht Heidelberg London 2007. 893 s. ISBN 978-1-4419-1749-2.
- [2] Corke, Peter: Robotics, Vision and Control. Springer-Verlag Berlin Heidelberg 2011. 570 s. ISBN 978-3-642-20143-1.
- [3] Corke, Peter: Robotics Toolbox for MATLAB, Release 9. Available on the Internet: <<http://www.petercorke.com/RTB/robot.pdf>>.
- [4] Mathworks MATLAB: R2012a Documentation [online]. Available on the Internet: <<http://www.mathworks.com/help/techdoc/>>.
- [5] Robotis Dynamixel Documentation [online]: Available on the Internet: <<http://support.robotis.com/en/>>.
- [6] VACEK, M.: "Optický bezpečnostný systém robotického pracoviska", Technical university of Košice, Faculty of Electrical Engineering and Informatics, Postgraduate Thesis, 2013.

# Využitie integrácie formálnych metód pri návrhu a analýze diskretných systémov

*Slavomír ŠIMONĀK*

Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika

slavomir.simonak@tuke.sk

**Abstrakt** – Témou príspevku je motivácia, prehľad vybraných riešení a prezentácia niektorých pôvodných výsledkov z oblasti integrácie formálnych metód. Prezentované pôvodné výsledky sa týkajú transformácie procesných algebier a Petriho sietí. Vytvorené transformačné nástroje boli prakticky využité pri analýze komunikačných protokolov.

**Kľúčové slová** – integrácia formálnych metód, Petriho siete, procesná algebra

## I. ÚVOD

Niektoré chyby pri návrhu systému môžu mať za následok obrovské straty. Hlavnou úlohou metód návrhu systémov je preto umožniť vývojárom konštruovať spoľahlivo pracujúce systémy bez ohľadu na ich zložitosť. Použitie formálnych metód (FM) pre návrh systémov je jednou z možností dosiahnutia uvedeného cieľa. Formálne metódy sú matematické techniky pre špecifikáciu a verifikáciu systémov. Samotné ich použitie však negarantuje korektnosť systému (nakoľko pracujú s formálnymi modelmi systémov, ktoré môžu byť nepresné či nekompletné), ale výrazne zvyšuje mieru porozumenia požiadavkám, ktoré sú na systém kladené [18]. Potreba použitia formálnej špecifikácie pri návrhu systému narastá so zložitosťou navrhovaného systému, ale závisí aj od oblasti jeho nasadenia. Pri vývoji systému bez použitia formálnych metód, jedinou možnosťou overenia funkčnosti systému vzhľadom na požiadavky je testovanie. Pri dostatočne zložitých systémoch testovanie môže preukázať ich chyby, nie však ich bezchybnosť [19].

Rozdielne formálne metódy poskytujú odlišné pohľady na navrhovaný systém a umožňujú využitie príslušného modelovacieho a analytického aparátu. Podľa tvrdenia, všeobecne uznávaného komunitou pracujúcou v oblasti formálnych metód, neexistuje jediná univerzálna metóda, ktorá uspokojivo zahŕňa všetky vlastnosti navrhovaných systémov [23], [20]. Možným riešením tohto problému je kombinácia niekoľkých metód [22], ktoré poskytujú komplementárne pohľady na navrhovaný diskretný systém a tým umožňujú vytvoriť optimálne podmienky pre jeho vývoj [24], [25].

## II. INTEGRÁCIA FORMÁLNYCH METÓD

O význame integrácie formálnych metód vypovedá aj vznik úspešnej série konferencií IFM (Integrated Formal Methods), organizovanej od roku 1999 s periódou 18 mesiacov. Konferencia je v súčasnosti pevnou súčasťou kalendára konferencií z oblasti formálnych metód a jej cieľom je poskytnúť priestor pre diskusiu o najvhodnejších spôsoboch návrhu hardvéru a softvéru za účelom zvýšenia ich spoľahlivosti a robustnosti. Obsahovo je možné konferencie IFM vnímať ako časť úsilia o zvládnutie komplexnosti (návrhu a analýzy systémov) kombináciou paradigiem špecifikácie a návrhu tak, že v rôznych etapách životného cyklu sú použité tie najvhodnejšie nástroje [1]. Čiastočne je to otázka kombinácie špecifikačných formalizmov (napríklad kombinácia stavovo-orientovaných a udalostne-orientovaných jazykov) pre vytvorenie integrovanej notácie, schopnej pokryť širšiu oblasť návrhového spektra. Veľmi významnou sa však stáva aj otázka verifikácie s využitím integrácie formálnych metód (napr. Forte - prostredie pre podporu priemyselnej verifikácie hardvéru používané spoločnosťou Intel [6] na báze integrácie model checkingu a dokazovania viet v rámci funkcionálneho programovacieho jazyka).

Integrácia formálnych metód zvyšuje ich význam ako intelektuálneho nástroja pri modelovaní a návrhu. Techniky integrácie zvyčajne zahŕňajú tieto postupy: vytváranie hybridných notácií, rozširovanie existujúcich notácií, prevody medzi notáciami, prípadne získanie širšej perspektívy

inovatívnym použitím existujúcej notácie. Aplikácia formálnych metód na komplexné systémy môže priniesť modelovanie rôznych aspektov a pohľadov, ktoré sú často vyjadrené pomocou rôznych paradigiem. Pokiaľ zložitosť je daná hlavne kvantitatívnymi prvkami (napr. veľkosť stavového priestoru), podporné nástroje urobili v poslednom období veľký pokrok. Ak ale zložitosť je zapríčinená potrebou kombinácie rôznych modelovacích náhľadov, nástroje sú často príliš doménovo-spcifické a teda nie veľmi vhodné pre konceptuálnu kombináciu. Integrácia by však nemala byť v prvom rade otázkou nástrojov, ale pozornej voľby konceptov, abstrakcií a formalizmov predtým, než začneme uvažovať o nástrojoch [26].

### III. PREHLAD VYBRANÝCH RIEŠENÍ V OBLASTI INTEGRÁCIE FM

V súčasnosti existuje množstvo riešení s využitím integrácie formálnych metód za účelom získania výhod pri opise a/alebo analýze diskretných systémov. Spomedzi nich vyberáme niekoľko zaujímavých prístupov.

#### A. Algebraická sémantika pre Hierarchické Petriho siete

V práci [27] je autormi prezentovaná algebraická sémantika pre P/T siete - PTNA (Place/Transition-Net Algebra), založená na procesnej algebre ACP. Je preukázané, že pôvodná P/T sieť aj jej algebraická reprezentácia majú rovnaké operačné správanie. Na rozdiel od iných postupov, akcie v prezentovanej algebre nekorešponujú s realizáciou prechodov, ale so spotrebou a produkciou značiek v miestach siete. Uvedená teória umožňuje zisťovať, či dve Hierarchické P/T siete majú rovnaké pozorovateľné správanie.

Štandardné P/T siete nedisponujú vlastnosťou de/kompozície, dôležitou pre modulárny návrh. Niektoré iné teórie pre opis paralelných systémov ako CCS, CSP, ACP však túto vlastnosť majú, preto bolo už vykonaných niekoľko pokusov zameraných na integráciu Petriho sietí a procesnej algebry. Niektoré z nich sa zameriavajú na sieťovú sémantiku pre procesné algebry, iné algebraickú sémantiku pre Petriho siete. Všetky tieto postupy však majú spoločnú vlastnosť, ktorou je korešpondencia medzi akciami v algebrách a prechodmi v sieťach.

V použitej triede P/T sietí je množina miest rozdelená na množiny tzv. vývodov a interných miest, kde vývody predstavujú konektory k prostrediu. Interná štruktúra siete teda môže byť ukrytá (black box). P/T siete s vývodmi predstavujú jednoduché Hierarchické P/T siete, ktoré okrem spomínaných miest a prechodov môžu obsahovať podsiete, ktorých vývody sú spojené s internými miestami, alebo vývodmi Hierarchickej siete. Pozorovateľné správanie v tomto prípade predstavuje jeho projekciu na vývody. Dve P/T siete sú považované za ekvivalentné, ak ich pozorovateľné správanie je rovnaké. Ako operačná sémantika pre P/T siete aj procesnú algebru sú použité prechodové systémy, ktoré umožňujú jednoznačné porovnanie správania.

#### B. Reprezentácia uzavretých CCS systémov Petriho sieťami

Transformácia CCS kompozícií do formalizmu Petriho sietí je opísaná v práci [29]. Pri dodržaní určitých obmedzení, kladených na syntax CCS, je výsledná sieť konečná. Realizácia jej prechodov korešponduje s komunikáciou v CCS kompozícii. Transformácia sa ukázala ako prospešná pri rýchlej detekcii uviaznutí (deadlock) v systémoch špecifikovaných pomocou CCS. Podľa autora, sémantika klasických Petriho sietí nie je tak silná ako sémantika CCS [28], [31], tieto však môžu slúžiť na určité účely, pre ktoré táto skutočnosť nie je prekážkou. Autor využil Petriho siete pri analýze uzavretých systémov (kompozícia procesov systému a prostredia) [30].

Ekvivalencia medzi komunikáciou v uzavretom CCS systéme (špecifikovanom množinou výrazov  $E$ ) a realizáciou viditeľných prechodov v Petriho sietí vytvorenej pomocou transformácie  $T_n(E)$  umožňuje využitie techník z oblasti Petriho sietí pre analýzu daného systému. Jednou z takých techník je generovanie stromu dosiahnuteľnosti. List v takomto strome predstavuje uviaznutie. Výskyt uviaznutí však môže byť vylúčený len konštrukciou celého stromu. Transformácia  $T_n$ , ako aj konštrukcia stromu dosiahnuteľnosti bola implementovaná vo funkcionálnom jazyku GOFER (balík CCSDD).

#### C. Meta-metóda pre integráciu formálnych metód

Integrácia metód v tomto prípade zahŕňa definíciu vzťahov medzi rôznymi metódami, ktoré majú byť spoločne využívané pre riešenie problémov. Prezentovaná je systematická technika pre integráciu formálnych metód založená na heterogénnych notáciách [20], [22], [23].

Dôležitou motiváciou pre integráciu formálnych metód je skutočnosť, že jedna metóda pravdepodobne nebude postačovať pri vývoji systémov so stále zvyšujúcou sa zložitou. Navyše integrácia metód bola využitá a preukázateľne prospešná v rôznych podobách (napr. Rolls-Royce, BT, Westinghouse, Praxis). Kombinácia notácií je chápaná ako prvý krok ku kombinácii metód.

*Heterogénne notácie* sú užitočné z viacerých dôvodov (jednoduchšie špecifikačné jazyky, jednoduchšie špecifikácie, úspešné využitie v praxi). *Heterogénna báza* je množina notácií, prevodov medzi formalizmami a formalizáciami, ktorá poskytuje formálnu sémantiku špecifikácií zložených z dvoch alebo viacerých notácií. V práci [21] prezentovaná báza pozostáva z množiny jazykov a prevodov definovaných medzi nimi (predikáty, Z, CSP, Larch a ďalšie notácie).

Špecifikácia zapísaná pomocou kombinácie notácií má potom sémantiku vyjadrenú pomocou jednej z formálnych notácií bázy. Existencia heterogénnej bázy umožňuje priradenie formálnej sémantiky kompozíciám parciálnych špecifikácií zapísaných v notáciách bázy. Neposkytuje však žiadne garancie týkajúce sa vhodnosti (prípustnosti) použitia špecifických notácií v kompozícii. Táto kompatibilita musí byť preskúmaná osobitne pre konkrétne prípady. Heterogénne notácie nie sú schopné riešiť všetky problémy integrácie metód (kompatibilita metód), avšak predstavujú systematický základ pre integráciu formálnych metód.

#### D. Milnerov CCS a Petriho siete

V práci [32] autori (Cindio, Michelis, Pomello, Simone) konštatujú, že CCS definuje podtriedu súbežných systémov definovateľných pomocou teórie Petriho sietí, konkrétne triedu systémov zloženú zo sekvenčných automatov so vzájomnými interakciami. Dôkaz je založený na konštrukcii izomorfizmu medzi podtriedou CCS (SCA-fna, System of Communicating Agents with fixed number of agents) a podtriedou Petriho sietí ( $SA_{PT}^2$  - podtrieda SA sietí). Podľa klasickej definície CCS modeluje systém ako kompozíciu ľubovoľnej množiny subsystémov (agentov), ktorých interakcia predstavuje vykonávanie párov akcií, považovaných za simultánne. Naopak SA siete modelujú systém ako kompozíciu pevného počtu subsystémov (procesov), ktorých interakcia prebieha vykonávaním ľubovoľného počtu akcií (prechodov), z čoho každá sa týka iného procesu. Porovnanie je založené na definícii podtried CCS a SA sietí a ukazuje, že majú rovnakú modelovaciu silu. Práca v závere obsahuje návrhy rozšírenia uvedeného izomorfizmu na celú triedu CCS.

#### E. Integrácia špecifikačných jazykov pomocou viacerých náhľadov

V práci [3] je prezentovaný všeobecný integračný mechanizmus pre špecifikačné jazyky, motivovaný čiastočnou špecifikáciou. Pre formalizáciu špecifikačných jazykov a definíciu relačného sémantického rámca je použitá teória kategórií. Systém je špecifikovaný pomocou viacerých náhľadov (viewpoint), pričom každý náhľad môže byť vytvorený pomocou iného špecifikačného jazyka, t.j. opis celého systému je heterogénny. Ako príklad je uvedený systém z oblasti telekomunikácií, špecifikovaný pomocou troch rôznych náhľadov. *Dátový náhľad* je špecifikovaný pomocou algebraickej špecifikácie (ADT), *stavový a operačný náhľad* pomocou OCL notácie a nakoniec *procesný náhľad* je vytvorený pomocou jazyka procesnej algebry podobnej CCS.

#### F. Objektovo-orientované modelovanie s využitím HL Petriho sietí

Práca [4] je venovaná transformácii opisov paralelných systémov zapísaných v objektovo-orientovaných formalizmoch do Petriho sietí za účelom automatizovanej verifikácie získanej Petriho siete s využitím existujúcich nástrojov pre analýzu dosiahnuteľnosti. V práci je prezentované rozšírenie jazyka  $B(PN)^2$  pre opis paralelných systémov o koncepty objektovo-orientovaného programovania. Rozšírenie má názov BOON (Basic Object-Oriented Notation) a umožňuje využívanie tried, dedenia, polymorfizmu a voľného viazania. Sémantika  $B(PN)^2$  je definovaná pomocou tzv. M-sietí (trieda vysokoúrovňových Petriho sietí). Podpora pre  $B(PN)^2$  a M-siete je implementovaná v rámci nástroja PEP toolkit, ktorý umožňuje simuláciu modelovaného systému a verifikáciu jeho vlastností pomocou model-checkingu.

#### G. Integrácia CSP-OZ s UML a Javou

Integrovaná formálna metóda CSP-OZ (CSP+Object-Z) v prepojení so štandardnými jazykmi pre návrh softvérových systémov (UML a Java) je prezentovaná v práci [5]. Cieľom je generovanie značnej časti CSP-OZ špecifikácie z UML modelu a následná transformácia formálnej špecifikácie do tvrdení (assertions) zapísaných v jazyku JML (Java Modelling Language). CSP-OZ špecifikácia slúži pre verifikáciu korektnosti UML modelu a JML tvrdenia zabezpečujú pridriavanie sa Java implementácie týchto formálnych požiadaviek počas behu programu. JML formou anotácií Java programov špeciálnym typom komentárov umožňuje využitie pre- a post-podmienok metód, invariantov a špeciálnych modelových premenných pre opis abstraktného stavu inštancií daného typu.



#### IV. VÝSLEDKY V OBLASTI INTEGRÁCIE PETRIHO SIETÍ A PROCESNEJ ALGEBRY

V nasledujúcom texte sú v stručnej podobe prezentované niektoré z dosiahnutých výsledkov v oblasti integrácie formálnych metód. Jedná sa o transformácie špecifikácií Petriho sietí [2] a procesnej algebry [10], [11].

##### A. Transformácia termov procesnej algebry ACP do jazyka Petriho sietí

Výsledky dosiahnuté autorom v oblasti transformácie termov procesnej algebry ACP do jazyka Petriho sietí boli publikované napr. v prácach [33], [34], [17]. Tieto výsledky zahŕňajú teoretický opis transformácie pre získanie sieťovej sémantiky termov procesnej algebry, návrh algoritmu ACP2Petri a jeho implementáciu v jazyku Java. Použitá je procesná algebra  $ACP_\epsilon$  (ACP rozšírená o pojem prázdneho procesu  $\epsilon$ ) a Petriho siete s definovanou množinou počiatkových a koncových miest. Jednotlivým akciám, z ktorých sú vytvorené algebraické termy, zodpovedajú tzv. elementárne siete. Z uvedených sietí, použitím definovaných operácií je konštruovaná výsledná sieť, zodpovedajúca zdrojovému procesnému výrazu (termu). Výslednej sieti je priradená interpretácia tým, že jednotlivé prechody sú označené menami akcií, z ktorých je zdrojový term vytvorený.

Uvedený postup zachytáva navrhnutý algoritmus (ACP2Petri). Najprv sú atomickým akciám procesného termu priradené zodpovedajúce elementárne siete, ktoré sú následne komponované pomocou operácií korešpondujúcich s operáciami procesnej algebry  $ACP_\epsilon$ , za účelom vytvorenia zodpovedajúcej Petriho siete. Na vytvorenie sietí je následne aplikovaná optimalizácia zjednodušujúca jej štruktúru vynechaním určitých prechodov s návěstím prázdnej akcie  $\epsilon$  a odstránením vzniknutých izolovaných častí. V práci [33] bolo preukázané zhodné lingvistické správanie sa zdrojového procesného termu i výslednej Petriho siete.

V rámci implementácie aplikácie ACP2Petri bol navrhnutý nový jazyk PAML (Process Algebra Markup Language), založený na technológii XML, pre zápis vstupnej špecifikácie ACP vo forme vhodnej na strojové spracovanie [16], [15]. Nástroj vytvára výstupný súbor (zodpovedajúca Petriho sieť) vo formáte PNML (Petri Net Markup Language) [35], ktorý je štandardným výmenným formátom pre Petriho siete, založeným na XML.

##### B. Transformácia Petriho siete na špecifikáciu v jazyku procesnej algebry APC

Pôvodné výsledky dosiahnuté v danej oblasti sú reprezentované definíciou algebraickej sémantiky pre Petriho siete a boli publikované v prácach [33], [9]. Prezentovaný postup využíva algebru APC, definovanú autorom za účelom opisu procesov reprezentovaných triedou bežných Petriho sietí (Ordinary Petri nets). APC koncepčne vychádza z procesnej algebry ACP, ktorá bola modifikovaná takým spôsobom, aby umožňovala pohodlnejší opis vyššie spomenutej triedy procesov.

Bol odstránený operátor komunikácie medzi procesmi ( $\mid$ ), ako aj s ním asociovaná komunikačná funkcia  $\gamma$ , nakoľko sa taká komunikácia u spomínanej triedy Petriho sietí nevyskytuje. Zavedený bol ale nový typ procesu - *procesný komponent*, ktorý nie je plnohodnotným procesom, čo sa prejavuje v jeho neschopnosti samostatne vykonávať akcie. Procesný komponent má iba schopnosť vzájomného spojenia sa s príslušným komponentom (komponentmi) za účelom vytvorenia procesu, schopného samostatného vykonávania jeho akcií. Procesy tohto typu boli zavedené za účelom prehľadnejšieho modelovania synchronizácie procesov, ktorá v prípade Petriho sietí je modelovaná prirodzene (prechodmi s viacerými pre-miestami). Ďalšou modifikáciou (voči ACP) je zavedenie kompozičnej funkcie  $\pi$  a binárneho operátora kompozície procesných komponentov ( $\mid\mid$ ). Modifikácie sa dotkli aj tabuľky prechodov (vplyvom zavedenia nového typu procesu) a systému axiém.

Špecifikácia v jazyku algebry APC je vytváraná tak, že každému miestu Petriho siete je priradená premenná definujúca množinu výpočtov danej siete iniciovaných v tomto mieste. Tieto premenné sú v práci označované ako E-premenné a ich hodnoty sú závislé na štruktúre zdrojovej Petriho siete. Preukázané je zhodné operačné správanie sa Petriho siete a jej algebraickej reprezentácie.

Implementácia uvedeného postupu v podobe nástroja Petri2APC [37] bola vytvorená v jazyku Java s využitím knižnice PNML Framework [36], určenej pre integráciu podpory jazyka PNML do nástrojov pre Petriho siete. PNML Framework poskytuje komplexné API pre načítavanie a ukladanie Petriho sietí vo formáte PNML. Výstupným formátom nástroja je PAML, ktorý bol rozšírený o odporu jazyka APC.

#### V. VYUŽITIE TRANSFORMÁCIE ACP2PETRI PRI MODELOVANÍ A ANALÝZE KOMUNIKAČNÝCH PROTOKOLOV

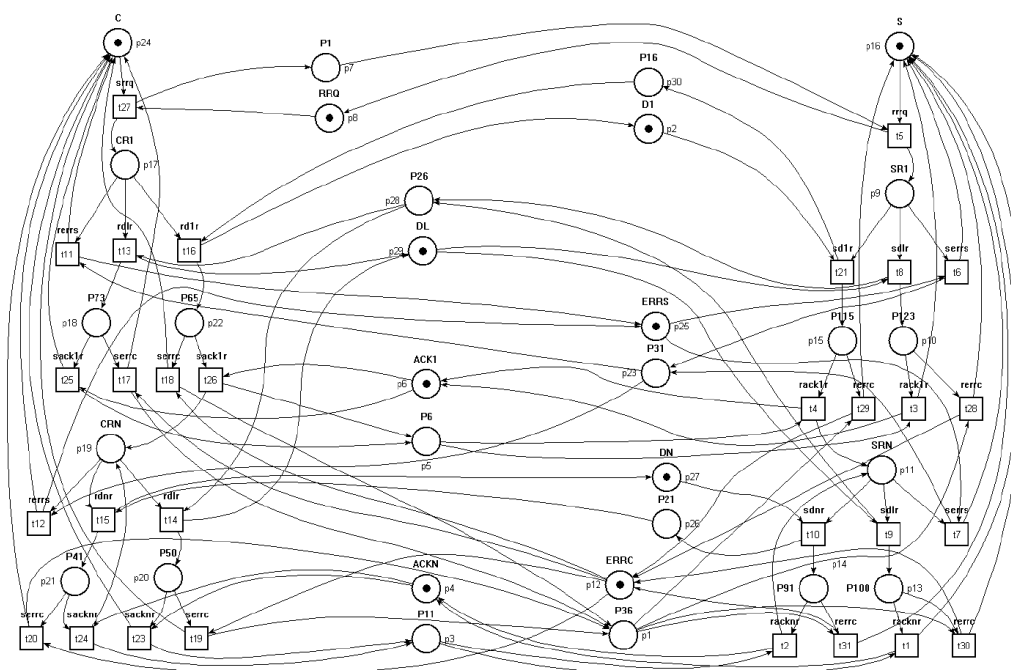
Komunikačný protokol definuje množinu pravidiel potrebných pre výmenu správ medzi komunikujúcimi entitami [12]. Množstvo sieťových a distribuovaných systémov [11], využívajúcich

komunikačné protokoly je v súčasnosti používaných v rámci aplikácií citlivých na bezpečnosť (safety-critical systems), kde dôvera v korektnosť komunikačného protokolu je nevyhnutnosťou. V prácach [7], [8] boli prezentované naše riešenia pre modelovanie a analýzu protokolov, založené na integrácii formálnych metód. Petriho siete a procesná algebra sú použité spoločne za účelom kombinácie výhodných vlastností oboch metód - tak prostriedkov procesnej algebry pre podporu pohodlného modelovania systémov, ako aj výnimočné analytické vlastnosti Petriho sietí.

Podľa našich skúseností je jednoduchšie vytvárať špecifikáciu komunikačného protokolu s využitím procesnej algebry, nakoľko de/kompozíciu a komunikáciu možno vyjadriť prirodzenejšie pomocou algebraických konštrukcií ako pomocou Petriho sietí. V prípade protokolov, systém zvyčajne pozostáva z komunikujúcich entít, komunikačného média, prípadne ďalších častí, ktoré možno špecifikovať osobitne a následne spojiť pomocou operácie paralelnej kompozície.

Na modeloch reprezentovaných pomocou Petriho sietí je možné vyšetrovať tak behaviorálne, ako aj štrukturálne vlastnosti systémov, pričom obidve možnosti prinášajú výsledky cenné pre analýzu komunikačných protokolov. Ústredným prvkom použitej metódy je automatická, sémantiku zachovávajúca transformácia algebraickej špecifikácie komunikačného protokolu na zodpovedajúcu Petriho sieť (ACP2Petri). Po realizácii transformácie je možné využiť výkonný analytický aparát Petriho sietí a tak objaviť prípadné defekty v konzistencii a korektnosti špecifikácie.

Navrhnutá metóda bola využitá pri verifikácii protokolov TFTP (Trivial File Transfer Protocol) [7] a ABP (Alternating Bit Protocol) [8]. TFTP je jednoduchý protokol pre presuny súborov medzi počítačmi, navrhnutý za účelom jeho jednoduchej implementácie a tak neobsahuje niektoré vlastnosti regulárneho FTP. Podporuje iba čítanie a zápis súborov z/na vzdialený server. Nepodporuje výpisy obsahu katalógov ani autentifikáciu používateľov. Keď sa počas prenosu objaví chyba, prenos je zastavený, takže je potrebné znova nadviazať spojenie a začať prenos. Operácie čítania a zápisu súborov sú vzájomne nezávislé, takže systém je možné dekomponovať a špecifikovať každú z operácií samostatne. Petriho sieť vytvorená na základe algebraickej špecifikácie operácie čítania súboru (TFTPR) protokolu TFTP je na Obr. 1.



Obr. 1 Petriho sieť pre operáciu čítania súboru (TFTPR) [7].

ABP predstavuje jednoduchý sieťový protokol pracujúci na spojovej vrstve OSI referenčného modelu. Je to protokol, ktorý zabezpečuje spoľahlivú komunikáciu cez médium, na ktorom môže dôjsť k strate správy. K dátam, ktoré odosielateľ posielajú príjemcovi sa pripojí informácia v podobe jedného bitu (0 alebo 1), ktorá slúži na potvrdenie úspešného prijatia správy. V prípade nezahyby kontrolných bitov proces odosielateľa nezačne odosielať novú správu, ale pokračuje v odosielaní poslednej odoslanej správy.

## VI. ZÁVER

V príspevku bol prezentovaný prístup k integrácii formálnych metód založený na transformácii procesných algebrií a Petriho sietí. Navrhnuté transformácie boli implementované v podobe

softvérových nástrojov [14]. Transformácia ACP2Petri bola prakticky využitá pri modelovaní a analýze komunikačných protokolov. Medzi perspektívne oblasti ďalšieho výskumu v danej oblasti patrí preskúmanie limitov vytvorených transformácií, ich doplnenie a rozšírenie. Podobne je možné uvažovať o návrhu transformácií ďalších rozšírených procesných algebier (CCS, CSP) do formalizmu Petriho sietí a ich implementácii v podobe transformačných nástrojov. Ďalším perspektívnym smerom je zahrnutie časového hľadiska do transformácií formálnych metód. Integrácia formálnych metód s časovým rozšírením by umožnila analýzu časovo-kritických systémov [13] uvedeným spôsobom.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (ITMS 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] Eerke A. Boiten, John Derrick, Graeme Smith (Eds.): *Integrated Formal Methods*, 4th International Conference, IFM 2004, Lecture Notes in Computer Science 2999, Springer-Verlag, 2004.
- [2] Desel, J., Juhás, G.: *What Is a Petri Net? Informal Answers for the Informed Reader*, volume 2128 of Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg, 2001.
- [3] Bujorianu, M., C.: *Integration of Specification Languages Using Viewpoints*, E. Boiten, J. Derrick, G. Smith (Eds.): IFM 2004, LNCS 2999, pp. 421–440, Springer-Verlag Berlin Heidelberg 2004.
- [4] Thanh, C., B., Kludel, H.: *Object-Oriented Modelling with High-Level Modular Petri Nets*, E. Boiten, J. Derrick, G. Smith (Eds.): IFM 2004, LNCS 2999, pp. 287–306, Springer-Verlag Berlin Heidelberg 2004.
- [5] Möller, M., Olderog, E., R., Rasch, H., Wehrheim, H.: *Linking CSP-OZ with UML and Java: A Case Study*, E. Boiten, J. Derrick, G. Smith (Eds.): IFM 2004, LNCS 2999, pp. 267–286, Springer-Verlag Berlin Heidelberg 2004.
- [6] Melham, T.: *Integrating Model Checking and Theorem Proving in a Reflective Functional Language*, E. Boiten, J. Derrick, G. Smith (Eds.): IFM 2004, LNCS 2999, pp. 36–39, Springer-Verlag Berlin Heidelberg 2004.
- [7] Šimoňák, S.: *Verification of Communication Protocols Based on Formal Methods Integration*, Acta Polytechnica Hungarica, Vol. 9, No. 4, 2012, pp. 117-128.
- [8] Šimoňák, S., Hudák Š., Korečko Š.: *Protocol Specification and Verification Using Process Algebra and Petri Nets*, Proceedings of CSSim 2009, Brno, Czech Republic, IEEE Computer Society, pp. 110-114, September 2009.
- [9] Šimoňák, S., Hudák Š., Korečko Š.: *APC Semantics for Petri Nets*, Informatica, Volume 32, 3, 2008, pp. 253-260.
- [10] Fokkink, W.: *Introduction to Process Algebra*, Springer-Verlag, 2007.
- [11] Fokkink, W.: *Modelling Distributed Systems*, Springer-Verlag, 2011.
- [12] Mauw, S., Veltink, G. J.: *Algebraic Specification of Communication Protocols*, Cambridge University Press, 1993.
- [13] Hudák, Š., Korečko, Š., Šimoňák, S.: *Reachability Analysis of Time-Critical Systems*, Petri Nets: Applications, Vukovar, Croatia, In-Teh, 2010, pp. 253-280.
- [14] Šimoňák, S., Peňko, I.: *PATool – A Tool for Design and Analysis of Discrete Systems Using Process Algebras with FDT Integration Support*, Acta Electrotechnica et Informatica, Vol.10, No.1, 2010, pp. 59-67.
- [15] Šimoňák, S., Hudák, Š., Korečko, Š.: *ACP2PETRI: a tool for FDT integration support*, Proceedings of 8th International Conference on Engineering of Modern Electric Systems, Oradea, Felix Spa, Oradea, Romania, 2005, pp. 122-127.
- [16] Šimoňák, S., Hudák, Š.: *ACP2PETRI - a tool for formal method transformations*, Electronic Computers and Informatics - ECI 2004, Košice-Herľany, September 22-24, 2004, Košice, pp. 98-102.
- [17] Šimoňák, S., Hudák, Š.: *Sieťová sémantika termov procesnej algebry APC*, Acta Electrotechnica et Informatica, Vol.4, No.1, 2004, pp. 55-59.
- [18] Clarke, E.M., Wing, J.M.: *Formal Methods: State of the Art and Future Directions*, Technical Report CMU-CS-96-178, Carnegie Mellon University (CMU), 1996.
- [19] Dijkstra, E.W.: *A Discipline of Programming*, Prentice Hall, 1976.
- [20] Paige, R.F.: *Formal Method Integration via Heterogeneous Notations*, PhD Dissertation, November 1997.
- [21] Paige, R.F.: *A Meta-Method for Formal Method Integration*, Proc. Formal Methods Europe '97, Lecture Notes in Computer Science, Springer-Verlag, 1997.
- [22] Paige, R.F.: *Future Direction in Program Transformations*, ACM SIGPLAN Notices, pp. 94-98, Vol. 32, No. 1, January 1997.
- [23] Paige, R.F.: *Heterogeneous notations for pure formal method integration*, Formal Aspects of Computing, 1999.
- [24] Butler, M.: *csp2B: A Practical Approach To Combining CSP and B*, Declarative Systems and Software Engineering Group, Technical Report DSSE-TR-99-2, 1999.
- [25] Hartel, P., Butler, M., Currie, A., Henderson, P., Leuschel, M., Martin, A., Smith, A., Ultes-Nitsche, U., Walters, B.: *Questions and Answers about Ten Formal Methods*, Declarative Systems and Software Engineering Group, Technical Report DSSE-TR-99-1, 1999.
- [26] Boute, R.: *Integrating Formal Methods by Unifying Abstractions*, Integrated Formal Methods, 4th International Conference, IFM 2004, LNCS 2999, Springer-Verlag, 2004.
- [27] Basten, T., Voorhoeve, M.: *An Algebraic Semantics for Hierarchical P/T Nets*, Computing Science Report, Eindhoven University of Technology, pp.32, 1995.
- [28] Olderog, E.R.: *Nets, Terms and Formulas*, Cambridge University Press, 1991.
- [29] Olszewski, J.: *Representing Closed CCS Systems by Petri Nets*, UNSW-CSE-TR-9412, University of New South Wales, Australia, pp.19, 1994.
- [30] Olszewski, J.: *Non-Interleaving Semantics for CCS and Fast Deadlock Detection*, School of Computer Science and Engineering, University of New South Wales, SCS & E Report 9317, 1993.
- [31] Goltz, U.: *On Representing CCS Programs by Finite Petri Nets*, LNCS 324, pp. 339-350, Springer-Verlag, 1988.
- [32] De Cindio, F., De Michelis, G., Pomello, L., Simone, C.: *Milner's Communicating Systems and Petri Nets*, selected

- papers from the 3rd European Workshop on Application and Theory of Petri Nets, Springer Informatik Fachberichte 66, 1983.
- [33] Šimoňák, S.: *Integrácia formálnych metód s využitím transformácií Petriho sietí a procesných algebier*, Dizertačná práca, Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2003, 112 p.
  - [34] Šimoňák, S.: *Formal Methods Transformation Optimizations within the ACP2PETRI Tool*, Acta Electrotechnica et Informatica, Vol.6, No.1, 2006, pp. 75-80.
  - [35] PNML.org - Petri Net Markup Language. <http://www.pnml.org/>
  - [36] Hillah, L., Kordon, F., Petrucci L., Trèves, N.: *Model engineering on Petri Nets for ISO/IEC 15909-2: API Framework for Petri Net types metamodels*, Petri Net Newsletter, (69):22-40, October 2005.
  - [37] Šimoňák, S., Hudák, Š., Korečko, Š.: *PETRI2APC: towards unifying Petri nets with other formal methods*, Proceedings of the Seventh International Scientific Conference Electronic Computers and Informatics ECI 2006, Košice - Herľany, Slovakia, Viena press, 2006.

# Secure Server Prototype based on Micro Kernel

<sup>1</sup>Viliam TOKARČIK, <sup>2</sup>František JAKAB, <sup>3</sup>Ivan KLIMEK, <sup>4</sup>Jaroslav LÁMER

<sup>1,2,3,4</sup> Department of Informatics, Faculty of Electrical Engineering and Informatics,  
Technical University of Košice, Slovak Republic

<sup>1</sup>viliam.tokarcik@studen.tuke.sk, <sup>2</sup>frantisek.jakab@tuke.sk, <sup>3</sup>ivan.klimek@tuke.sk,  
<sup>4</sup>jaroslav.lamer@tuke.sk

**Abstract** — This paper deals with the examination and describes creation of secure server which uses the microkernel. Microkernel has different implementation than other types of kernels. It can be hybrid or monolithic. Microkernel is divided into modules that are separated and independent. When an error occurs in one of the modules, the whole system is isolated from this error. From this security view, microkernel is more interesting than other kernels in operating system.

**Keywords** — secure server, prototype, microkernel, communication, operating system

## I. INTRODUCTION

In paper are examined the vulnerabilities of the computer environment without or with low possibilities to attack computer system. Computer systems have to be physically protected from the outside as well as from the inside – this is achieved using secured software, security policy and appropriate strategies.

Solutions introduced in this project are focused on internal security. Except for need to increase security of information system there is also requirement to deal with security of the whole environment, i.e. environment where this system operates. Bearing in mind microkernel and its limited functionality, solution has the capability to provide protection against intrusion into the kernel operating system. Another advantage of microkernel is limitation of source code in contrast to other kernel types, e.g. monolithic and hybrid. From the perspective of influence by programming errors security of microkernel is kept in way that its low size and implementation using simple language helps and simplifies the verification of the source code.

## II. GOALS

Specific focus is put on the implementation of secure server prototype using microkernel. Fulfilling of this goal includes wide variety of appropriate microkernels that are capable to communicate with Linux operating system. Main idea is to develop a mechanism utilizing which security of operation system is to be increased.

Research provides an insight into proper ways of how to implement and build a prototype server. Last but not least elimination of insufficiency of known solutions is presented.

## III. ANALYSIS

The choice of kernel operating system is influenced by variety of kernels and differences among them. Subchapter A describes differences between kernels.

### A. Kernel architectures

Kernels in terms of division to modules are analyzed and according to architecture are divided into [1]:

- monolithic kernel,
- microkernel,
- hybrid kernel.

**Monolithic kernel** – in monolithic kernel, all operating system services run along the main kernel process as seen in Fig. 1. This approach ensures a full and powerful access to hardware. The major disadvantages of monolithic kernels are dependencies between system components. Error in drivers of device can cause a crash of the entire the system. Authors in [1] points out another fact – large kernels may be too difficult to maintain.

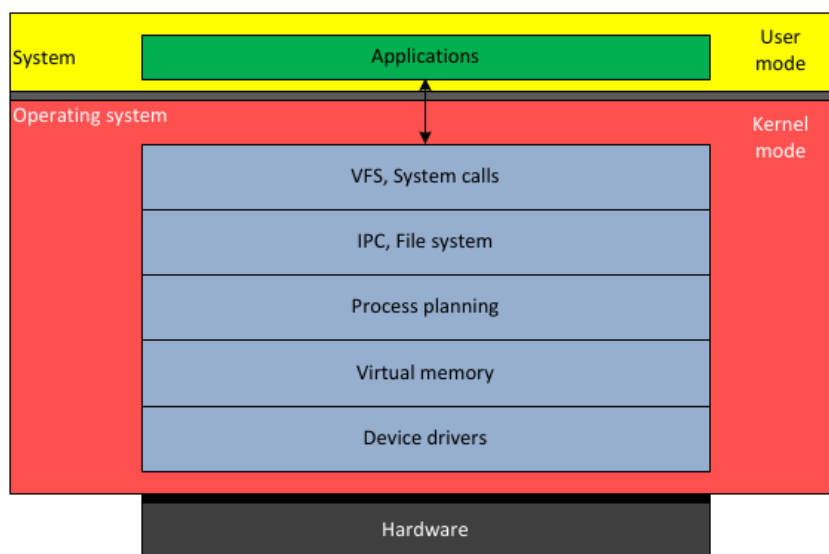


Fig. 1 Monolithic kernel (Kernel mode) in operating system structure

Monolithic type of kernels is mostly used in UNIX operating systems. It contains basic functionality, device drivers (driver is a small program that allows the operating system to interact with the hardware, such as disk drives, network cards etc.). This is a traditional design of UNIX systems. Monolithic kernel is a program that contains everything what is needed to execute each kernel task. Each element that should be available to more programs and cannot be located in the library, is in the kernel space. As stated by [1] [2], these parts include: device drivers, scheduler, memory management, file systems, and network stack.

**Microkernel** – also known as uK is the lowest level of software abstraction aimed at hardware. Architecture of microkernel is depicted in Fig. 2 and contains only core functions. Its responsibility includes managing of system resources, connections of applications and data processing on hardware level [3] [4] [5].

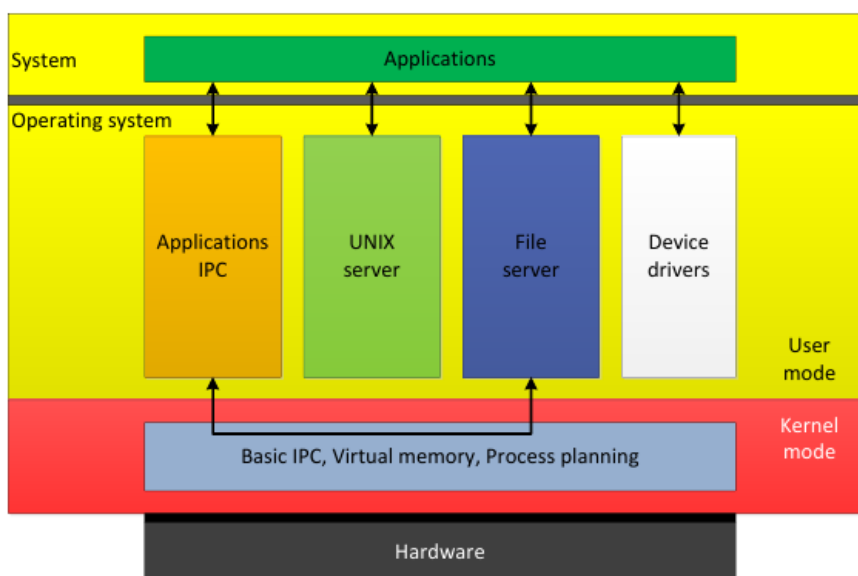


Fig. 2 Microkernel (Kernel mode) in operating system structure

Microkernel in opposite to monolithic kernel has enhanced functionality, e.g. networking layer, device drivers, file systems and so on. This functionality is implemented outside of

privileged mode in user environment. These services are running as servers and communicate with each other via IPC (Inter Process Communication). Using this architecture and minimization of operations, increase of security is implemented. Kernel is less likely to fail because of absence of redundant code [3] [4] [5].

**Hybrid kernel** – are similar to microkernels however this type of kernels include code to increase performance. They represent the compromise implemented even before the demonstration that microkernels provide higher performance. Hybrid kernels are extended by some properties of monolithic kernels. These kernels cannot load modules in runtime and contain code which executes in kernel space to run faster than as a part of user environment. Hybrid kernels may be seen as a compromise between monolithic kernels and microkernels – it means running of some services like networking and file system services in kernel space for performance reduction of microkernels while kernel code, e.g. device drivers, executes in users space [6] [7].

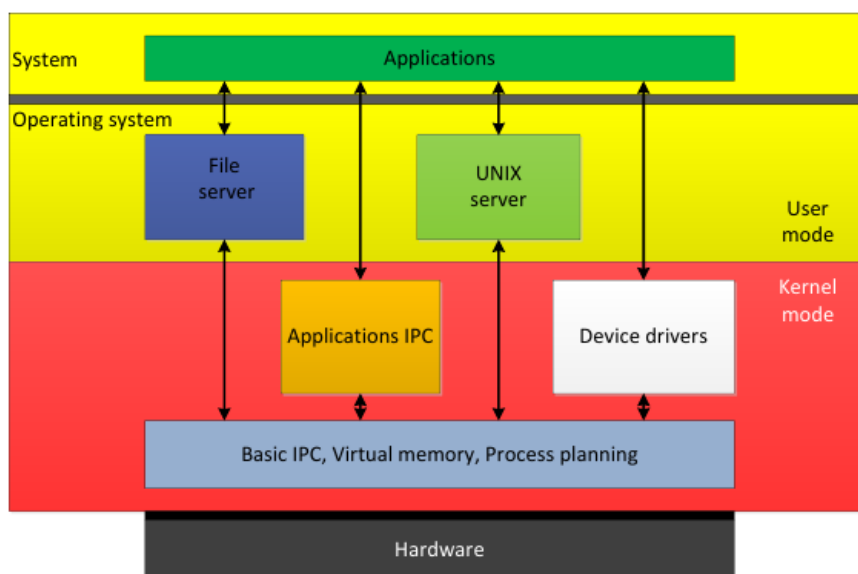


Fig. 3 Hybrid kernel (Kernel mode) in operating system structure

#### IV. SOLUTION AND RESULTS

##### A. Analysis of Utilized Microkernel seL4

SeL4 was chosen as a microkernel to be utilized in this project since its safety should be better when compared to other microkernels due to the fact that it was formally verified [8]. Another advantage of this microkernel is support of Wombat operating system. This operating system is designed to be portable between different processors architectures, the hardware platforms and versions of the Linux kernel. Emphasis during design was placed on acceptable price when adding new architectures and further on updates of the Linux kernel [9].

Collaborations of this microkernel and operating system Linux is implemented like new processor architecture, *L4*. New directories *arch/l4/* and *include/asm-l4/* were added to source code tree [9]. Application that is running directly in the microkernel does not provide Linux functions and also does not have possibility to collaborate with native applications running under Linux. Just mentioned was the problem we focused on and successfully solved. Solution consisted from creating a communication channel between the microkernel application and Linux.

**Communication channel** – consists of microkernel part and program in Linux. Special thread is created in Linux which provides data that are sent to program running in microkernel and vice versa.

**Communication channel component** – this was created in the catalog structure and placed in: *code/source/components*. In this place in the component catalog is source code, header files and configuration file for the compiler *Makefile*. Component of the communication channel is automatically initialized and launched at system startup.

**Linux part of communication channel** – contains source code of communication channel on Linux kernel side is in directory *wombat/kernel/arch/14/kernel*. This code was placed here because of (middle-layer) *L4* architecture implementation, further it uses calls of *seL4* microkernel and utilizes calls of microkernel as well.

Own kernel thread was created to secure always-ready state of communication channel. At the first was created kernel thread in main loop of kernel. Later was found that it is not appropriate to initialize channel in this place. Main problem was that Linux kernel has not yet initialized all its functions and processes – IRQ handler of interruptions has not started hence it was impossible to work with time. Another part of the problem was absence of memory management and functions like *vmalloc*, *kmalloc* and *kfree* were not available.

Thread was initialized after start of Linux kernel. It means that kernel was in system state *SYSTEM RUNNING*. This state is acquired in function *init\_post(void)* in file *wombat/kernel/init/main.c*. After execution of our thread kernel executes console */dev/console* and starts initialization of operation system */sbin/init*. Presented problem was solved by execution of our thread in this place.

### B. Measured results

The solution was tested for system load and message sending speed. The results were transformed into graphs. Each line in graph represents measured load or speed values for custom priorities of solution components in the order – wombat distribution priority, component communication channel priority and communication channel process priority.

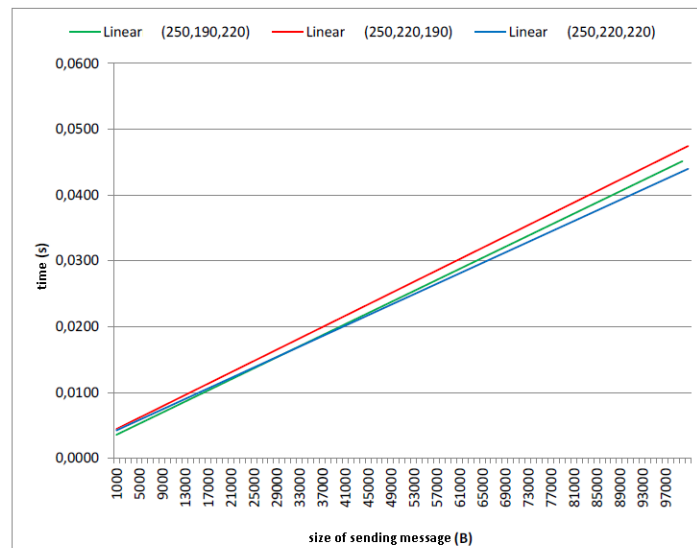


Fig. 4 Comparison of sending speeds - sending from component to Linux

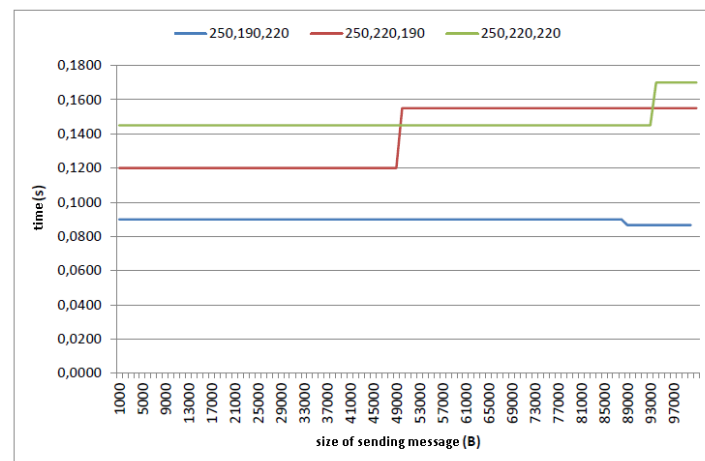


Fig. 5 Comparison of system load during message sending from component to Linux



## V. CONCLUSION

Safety server prototype which uses microkernel was created, where SeL4 is used as microkernel. This microkernel is formally verified and thus considered to be secure. Wombat was used as an operating system by server since it is based on Linux operating system and hence all Linux application can be used.

This article solves the communication problem between applications running on microkernel and applications running directly in Linux. Therefore the main addition of this article is the description of implementation of communication channel between described parts of system (applications).

This solution removing communication barrier between applications, they are directly running on microkernel and applications running on Linux or Linux functions.

## ACKNOWLEDGMENT

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development. Operational Program funded by the ERDF (20%).

## REFERENCES

- [1] ROCH, B.: Monolithic kernel vs. Microkernel. TU Wien, 2005.
- [2] JANSSEN, C.: Open Sources: Voices from the Open Source Revolution [online]. [cit. 2013.10.22] <<http://www.techopedia.com/definition/27003/monolithic-kernel>>.
- [3] MICHAEL, G. 1990: Micro-kernel Architecture Key to Modern Operating Systems Design [online]. [cit. 2013.10.22] <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.52.3352&rep=rep1&type=pdf>>.
- [4] UHLIG, V.: Scalability of Microkernel-Based Systems [online]. [cit. 2013.10.22] <[http://os.ibds.kit.edu/downloads/publ\\_2005\\_uhlig\\_scalability\\_phd-thesis.pdf](http://os.ibds.kit.edu/downloads/publ_2005_uhlig_scalability_phd-thesis.pdf)>.
- [5] LEE et al. 2008 Design of the L4 Microkernel Based Lightweight Mobile Middleware for Mobile Phone.
- [6] KANDUKURU, S. 2012: What is an Operating system? – Types of Kernel [online]. [cit. 2013.10.22] <<http://thelinuxdesk.com/tag/hybrid-kernel/>>.
- [7] GRIFFITHS et al. 2006: Hybrid vs. monolithic OS kernels: a benchmark comparison.
- [8] KLIEN et.al.: seL4: Formal verification of an OS kernel. In Proceedings of the 22<sup>nd</sup> ACM Symposium on Operating Systems Principles, USA, October 2009.
- [9] LESTIE et.al.: Wombat A Portable User-Mode Linux for Embedded Systems.

# Interactive 3D Visualization and Organization of Objects in Web Environment

<sup>1</sup>Tomáš ŽELEZNÍK, <sup>2</sup>Zdeněk HAVLICE, <sup>3</sup>Matúš KIRCHMAYER

<sup>1</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>2</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>3</sup>Matsuko s.r.o., Košice, Slovak Republic

<sup>1</sup>tomas.zeleznik@student.tuke.sk, <sup>2</sup>zdenek.havlice@tuke.sk,

<sup>3</sup>matus.kirchmayer@matsuko.com

**Abstract** – This work deals with the analysis of the technologies used during creating and implementing the interfaces of Web3D application called Webfield. Application is created by means of the scripted language JavaScript, cascade style CSS3 and markup language HTML5, it is built upon the design pattern Model-View-Controller. Part of the work is to experimentally rebuild the existing framework written in ActionScript language into Haxe language with the subsequent partial compilation into JavaScript language.

**Keywords** – Web3D application, scripting language JavaScript, cascade style CSS3, markup language HTML5, programming language Haxe

## I. INTRODUCTION

Browsing the web pages when the whole user interface is in 3D space, is not very common even today though it is enabled by hardware equipment for the owners of the devices. Such visual display of the web content is more interesting, more interactive and more modern for the users.

Nowadays most of the users uses tabs or panels while browsing, however, their usage in greater number is disarranged. The goal is implementing the application with three-dimensional user interface, better and more arranged organisation of pages (tabs) with the possibility to use it on mobile phones or desktops without installation.

Creation of the application called Webfield has consisted of

- analysis of the technologies for the implementation,
- finding the suitable IDE,
- design, by creating the UML diagrams,
- implementing and testing the prototype.

There are many web technologies for creating the content of the applications or pages with much better rendering quality, for instance Java3D, Unity3D and other [1]. Since the goal is to create spatially simple and multiplatform-supported application therefore it would not be suitable to use these technologies.

Part of the work is the experimental rebuild of existing framework, small part of which is linked with the created Webfield prototype. Usage of Haxe language and its subsequent compilation into JavaScript language is, regarding the performance, very similar to natively written JavaScript code. [2]

Work is partially related to Bachelor Thesis [3], in which implementation of the similar application by means of ActionScript technology, which is still not natively supported on mobile phones nowadays and is substituted by technologies used for the implementation in this work, is reasoned and described.

## II. GOALS

Main goal of the work is testing the web technologies HTML5, CSS3 and JavaScript, by creating the Web3D prototype, which are currently used for creating the common 2D web pages (applications) due to their suitability at creating 3D graphic interfaces. Another aim is to experimentally test the rebuild of the framework written in ActionScript language into Haxe language, with the subsequent compilation in to JavaScript language. [4]

### III. CHOICE OF THE TECHNOLOGIES FOR THE WEBFIELD APPLICATION DEVELOPMENT

Software tool for the application design by means of UML Enterprise Architect models has been chosen mainly due to its usage in the company as well as for its ability of reverse engineering of the source code into UML models of ActionScript language.

Integrated development environment IntelliJ IDEA has been, after the installation of the additional plug-in for the support of Haxe language, the best choice due to supporting the several languages (ActionScript, JavaScript, Haxe, HTML, CSS for our usage) and smart code completion.

ActionScript and HTML5 languages and CSS3 style, as opposed to other technologies (ActionScript, WebGL, Java3D, etc.), have been chosen mainly due to its greater support in the several platforms, mainly mobile phones without need to install the additional software. For running the application implemented by means of the chosen technologies, only browser of the web pages with newer update is needed. Choice of the Hammer.JS library for capturing the user input on the touch screen was, as opposed to other tested greater libraries such as Quo or jQuery Mobile, easier to implement.

Haxe language was used for the framework rebuild due to its possibility to compile into several known languages and static type control.

For creating the pop-up jQuery Mobile library has been used since its components are easy to install, possess the modern design and are supported on the mobile phones, as well.

### IV. DESIGN OF WEBFIELD APPLICATION INTERFACES

Design of application interfaces is divided into design of user interface and interface for server connection. Since prototype Webfield is a Web application, we use the most widely used design pattern for this type of application, namely the Model-View-Controller. Design of the respective parts of this pattern is as follows:

- Model - database and script running on the server is created by means of Google App Engine since this model is used from the existing similar application, it will not be described further in detail,
- View - user interface created by means HTML5, CSS3, JavaScript technologies,
- Controller - logical part of the application, interface for connecting the user interface with model created by means of JavaScript language.

From the graphic point of view the application will consist of the great number of tiles that may be manipulated variously by user (see the Figure 1).

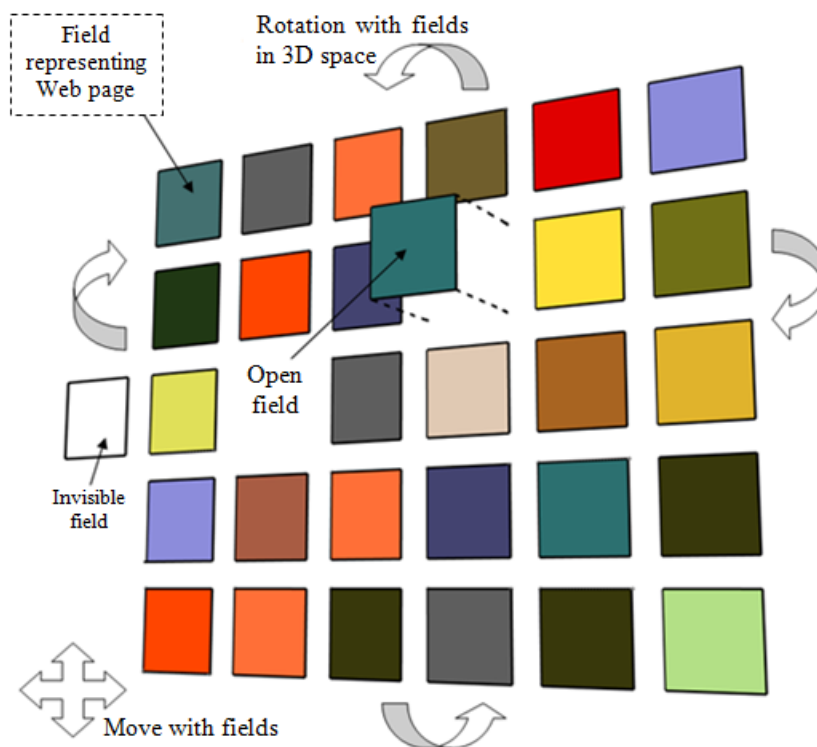


Fig. 1 Design of Webfield's GUI.

Communication between the application and server is carried by means of sending JSON (JavaScript Object Notation) objects with the usage of AJAX technology. Model of the application is built upon Google App Engine which is the platform by Google company used for web applications, it provides the space for the data on the server, database, it enables running the scripts (in our case Python) and manages the data communication between the application and Internet.

## V. REBUILD OF THE SOFTWARE FRAMEWORK

For testing purposes, it was needed to link the created interfaces by means of JavaScript language with the existing software framework written in ActionScript language. We decided to rebuild this software framework into Haxe language due to practical testing of this language usability and for the advantage of the later compilation into several languages. In the future it is also planned to rebuild the whole prototype into Haxe language.

Rebuild of the software framework consisted of finding the syntactic differences between the Haxe and ActionScript languages and subsequent gradual rebuild of the concrete packages and classes. Main syntactic differences between the languages are:

- basic data types,
- defining the class and constructor,
- typecasting,
- in cycle for and for each,
- usage of some specific functions such as explicit typecasting or finding the index of element in array,
- semicolon is mandatory after each line,
- in substitution of the data type Object of ActionScript language, it is not always possible (nor recommended) to use the similar type Dynamic of Haxe language but it is necessary to implement the concrete type such as Map, Array, etc.

Haxe language as well as ActionScript enables the usage of dynamic type, however, advantage of static typecasting control is lost. Using the dynamic type in Haxe language is useful for instance when using the third parties libraries or in specific cases when it is necessary for one variable to be of different types.

## VI. EXPERIMENTAL APPLICATION TESTING

Webfiled application has been tested on devices described in Figure 2.

On Asus laptop application was tested in the four most widely used browsers with the latest updates available at the time of writing this work. In Google Chrome and Opera there has been no problem in using them. Mozilla Firefox does not support the smooth rotation, scrolling, zooming the tiles in or out. Internet Explorer does not allow to use the application, although the application starts, the scene there is only one tile is displayed with which it is not possible to interact.

Both of the tested tablets of brands Apple and Asus allow full usage of application by using hand gestures.

The prototype has partial support in mobile phone HTC Desire X using the built-in browser, where it is not possible to move, rotate or resize the tiles by using the zoom in or out. By using the software Google Chrome usage of Webfield is again without any issues. Mobile phone Lumia 920 with Internet Explorer Mobile, like desktop version, does not support using the application (displays one tile, with which it is not possible to interact).

The last test was a platform e-book reader from Amazon - Kindle Paperwhite the second generation. Although this device has an electronic monochrome ink display and is not designed to use dynamic applications, it allows for partial support to use prototype Webfield. Implemented experimental browser does not support rotation of tiles and opening web pages associated with the tile, but allows them to edit or view any reference of the site that can then be manually opened in new tab after copying it.

It is important to say that the response rate, while using the application with the exception of hardware and software of the devices, has greatly depended on the amount of the created tiles. For the full usage of the tiles depending on the platform, the user should not create more than 100 tiles which is redeemable during the normal use. Final GUI of Webfield is depicted in Figure 3.

## ACKNOWLEDGMENT

This work is the result of the project implementation: Development of the Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120030) supported by the Research & Development Operational Program funded by the ERDF. This work was developed in MATSUKO s.r.o. company in Kosice.

Device	Operating system	Web browser	Support
Notebook ASUS N61jv	Windows 7 Professional 64-bit version, Service pack 1	Google Chrome 34.0	full
		Opera 20.0	full
		Mozilla Firefox 28.0	partial
		Internet Explorer 11.0	poor
Tablet Apple iPad 3.generation	iOS 7.1	Safari 7	full
Tablet ASUS ME301t	Android 4.2.1	Google Chrome 34.0	full
Cell phone HTC Desire X	Android 4.1.1	Integrated browser	partial
		Google Chrome 31.0	full
Cell phone Nokia Lumia 920	Windows Phone 8.0	Internet Explorer Mobile 10.0	poor
E-book reader Amazon Kindle Paperwhite 2.generation	Linux 5.4	Experimental	partial

Fig. 2 Support of Webfield on different platforms.

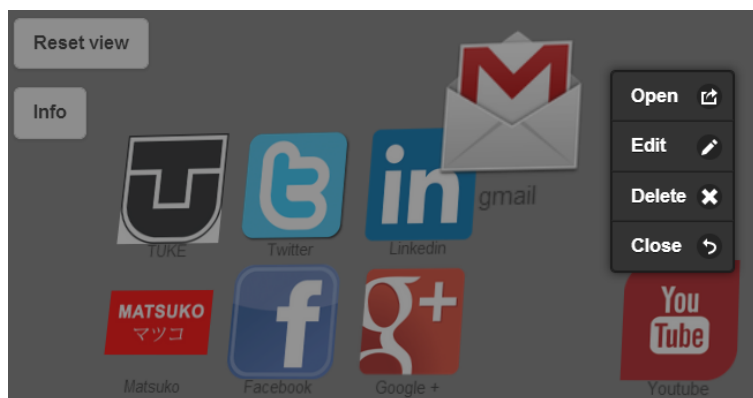


Fig. 3 Final GUI of Webfield.

## VII. CONCLUSION

By means of created user interface of Webfield application and its subsequent linking to server, we have testified that currently using the HTML5, CSS3 and JavaScript technologies is not suitable only for the development of two-dimensional web pages and applications, as it is in most of the cases, but also for the fully three-dimensional interactive content. We have created the functional prototype of three-dimensional application executable in the mobile phones with the touch screen and classic desktops without need to install any additional software. By testing the prototype in various devices we have verified that implementation by means of the used web technologies is sufficient for the smooth interaction of user with the application environment.

By rebuilding the software framework from ActionScript language into Haxe language we have found out that multiplatform Haxe language possesses a great potential while being implemented now or in the future. Although the main aim of the work was to create the prototype for testing the user interface, before the prototype is suitable for the full usage by user, it is necessary to program signing in and registration, as well.

For the future there are another possibilities to enhance the prototype, for instance to link it

with social networks or to create the better nested organisation of the tiles when after clicking one tile, subcategory of others as a tree hierarchy would open.

#### ACKNOWLEDGMENT

This work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development

#### REFERENCES

- [1] Yanxing Ou, Ruomei Wang, Guanwei Xiao: Design and Implementation of Virtual Prototype Apartment Based on Web3D. In: Digital Home (ICDH), 2012 4th International Conference on Digital Home, Danvers, MA, USA: IEEE Computer Society, Nov. 23-25, 2012, pp. 161-166.
- [2] Julien Richard-Foy, Olivier Barais, Jean-Marc Jézéquel: Efficient high-level abstractions for web programming. In: Proceedings of the 12th international conference on Generative programming: concepts & experiences, New York, NY, USA: ACM, 2013, pp. 53-60.
- [3] Tomáš Železník: Interactive 3D Visualization of the Web and Integration of Web3D Applications: Bachelor thesis. Košice: TU FEI, 2012.
- [4] Tomáš Železník: Interactive 3D Visualization and Organization of Objects in Web Environment: Diploma thesis. Košice: TU FEI, 2014.

# Nový optický vláknový refraktometrický systém ovládaný cez Web

<sup>1</sup>Tomáš CICANIČ, <sup>1</sup>Luboš OVSENÍK, <sup>1</sup>Ján TURÁN

<sup>1</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>lubos.ovsenik@tuke.sk, <sup>1</sup>jan.turan@tuke.sk

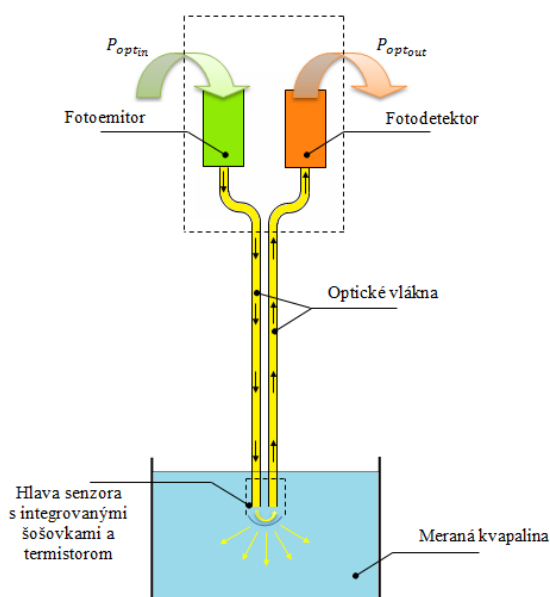
**Abstrakt** — Predložená práca sa zaoberá vzdialeným laboratórnym meracím systémom ovládaným cez Web pre výučbu aplikovanej fotoniky. Popisuje laboratórny optický vláknový refraktometrický systém používaný na meranie indexu lomu rôznych kvapalín.

**Kľúčové slová** — Web-ovládané meracie systémy, vzdialené meranie cez WWW, optický vláknový refraktometer, meranie indexu lomu kvapalín

## I. ÚVOD

V dnešnej dobe existuje veľké množstvo meracích pracovísk pre študentov a ich sprístupňovanie cez Web sa teší čoraz väčšiemu záujmu ako zo strany študentov, tak aj zo strany pedagógov. Cez Web-ovládané meracie pracovisko pomáha k prechodu od teoretických znalostí k praktickým skúsenostiam, ktoré by nepochybne mali byť súčasťou každého štúdia. Preto cieľom našej práce bol vývoj softvérového vybavenie pre unikátne meracie pracovisko s laboratórnym optickým vláknovým refraktometrom, ktoré by umožňovalo nielen jeho ovládanie cez Internet, ale aj spracovanie nameraných výsledkov.

Optický vláknový refraktometer nám umožňuje merať veľmi malé zmeny indexu lomu rôznych kvapalín, preto má veľké uplatnenie v mnohých odvetviach priemyslu (napr. potravinársky priemysel, chemický priemysel, atď.).



Obr. 1 Princíp meracieho systému indexu lomu kvapalín.

Aby sme mohli merať index lomu určitej kvapaliny, k tomu je potrebné použiť niektorý zo spôsobov na to určených. Naš merací systém využíva nasledujúci princíp.

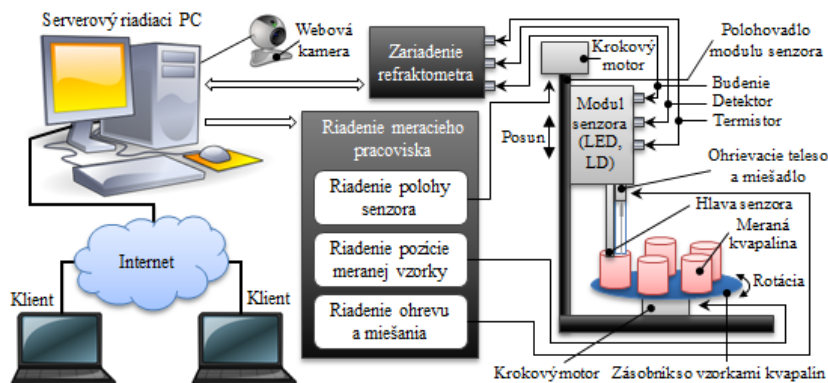
Kolmé osvetlenie jedného konca sklenenej tyčinky vzhľadom na jej os (obidva konce tejto tyčinky sú zbrúsené) spôsobí pri jej uložení vo vzduchu, kedy platí  $n_{\text{vzduch}} \ll n_{\text{tyčinka}}$ , v nej také šírenie svetla, že sa táto sklenená tyčinka chová ako vlnovod, tzn. žiarenie neprejde cez plášť von, pretože  $\theta_d > \theta_c$ . Takým vlnovodom bude aj optické vlákno (OV) zohnuté do tvaru písmena U. Kvapalina s indexom lomu rovnakým ako index lomu skla, do ktorej ponoríme sklenené koleno, vytvorí také podmienky, že svetelné žiarenie, ktoré predtým viedlo jadro optického vlákna, prejde bez lomu do nej, čoho výsledkom bude na výstupnom konci OV nulová intenzita žiarenia detegovaná na fotodetektore. V prípade kvapaliny s indexom lomu o dačo menším ako index lomu jadra OV, nastáva prechod časti svetla do kvapaliny a časť optického žiarenia sa odráža z rozhrania späť do jadra. Zväčšujúci rozdiel medzi indexom lomu jadra OV a indexom lomu obklopujúcej kvapaliny sa prejaví na detegovanej intenzite žiarenia, ktorá porastie. To sa využíva pri meraní indexu lomu kvapaliny ktorú skúmame, kedy jeho hodnotu získame z nameranej intenzity žiarenia prechádzajúceho vo vnútri jadra OV. Takýto merací systém indexu lomu je na Obr. 1.

## II. LABORATÓRNY OPTICKÝ VLÁKNOVÝ REFRAKTOMETRICKÝ SYSTÉM

### A. Web-ovládané meracie pracovisko

Web-ovládané meracie pracoviská poskytujú veľmi účinný spôsob vzdelávania. Takto sprístupnené pracovisko umožňuje pracovať 24 hod. denne, pre viaceré vzdialené skupiny záujemcov, pričom tieto môžu merania na tomto pracovisku vykonávať aj spoločne, čo si samozrejme vyžaduje ich komunikáciu navzájom a používanie výpočtových zdrojov.

Aby bola možná komunikácia s hardvérom meracích pracovísk cez Web, sú potrebné aplikácie, ktoré toto zabezpečia. Preto aj optický vláknový refraktometrický systém využíva na ovládanie cez Web na to určenú aplikáciu, ktorá pracuje so vzťahom klient-server, kde na jednej strane je klient (vzdialený používateľ) a na tej druhej server (serverový riadiaci PC). Štruktúra hardvérov tohto systému je ukázaná na Obr. 2.



Obr. 2 Štruktúra Web-ovládaného refraktometrického systému.

### B. Softvérová výbava ovládania optického vláknového refraktometrického systému cez Web

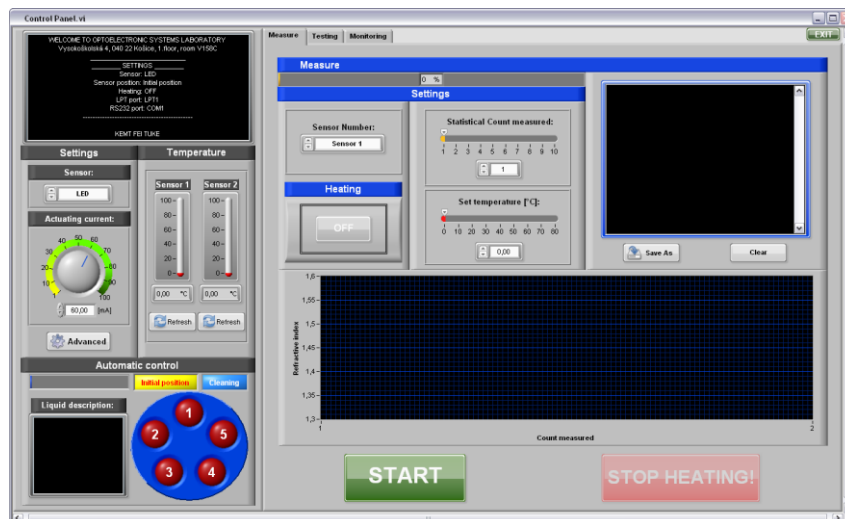
Softvérová výbava sprístupňujúca optický vláknový refraktometrický systém a jeho ovládanie aj cez Web sa skladá z troch častí. Prvá časť tvorí program vyvinutý v programovacom a vývojovom prostredí LabVIEW, ktorý na strane servera (serverový riadiaci PC) umožňuje ovládať laboratórny optický vláknový refraktometer pripojený cez sériové a paralelné rozhranie k tomuto serveru a ktorý ďalej pomocou implementovaných nástrojov sprístupňuje toto ovládanie vzdialenému používateľovi (klient) a to až do takej miery, ako by bol na meracom pracovisku. Druhá časť softvérovej výbavy obsahuje webovú stránku, ktorá je akousi vstupnou bránou pre vzdialeného klienta. Posledná tretia časť využívajúca Skype poskytuje vizuálnu komunikáciu prostredníctvom webovej kamery.

#### Ovládací program

Hlavnou činnosťou tohto programu je meranie indexu lomu, pričom k tomu je možné použiť jeden alebo dva senzory (LED alebo LD). Program umožňuje vykonávať meranie pomeru optických výkonov (pomer budenia ku detekcii), zisťovať teplotu skúmanej kvapaliny a spätne merať budiaci prúd. Nakoľko sa jedná o automatizované meracie pracovisko, je tu k dispozícii tiež ovládanie vertikálneho pohybu senzora a rotačného pohybu taniera obsahujúceho nádoby s



kvapalinami a taktiež riadenie zapínania/vypínania ohrievacieho telesa spolu s miešadlom. Na základe dostupných možností merania sú tak v programe implementované tri funkcie, a to Measure (Merač), Testing (Testovač) a Monitoring (Sledovač). Možnosť používať všetky tieto funkcie a nástroje na ovládanie meracieho pracoviska s laboratórnym optickým vláknovým refraktometrom je ako na strane serverového riadiaceho PC, ktorý je prepojený s týmto pracoviskom cez LPT a COM porty, tak aj zo strany vzdialeného na počítačovú sieť pripojeného používateľa (Obr. 3).



Obr. 3 Hlavné okno ovládacieho panela v programe Optical Fibre Refractometric System.

#### Webová stránka

Skôr než klient môže poslať žiadosť serverovému riadiacemu PC o získanie kontroly nad ovládacím panelom programu Optical Fibre Refractometric System, musí byť prihlásený na webovej stránke s adresou: <http://los.feit.tuke.sk/ofrs/>. Táto stránka umožňuje vzdialenému používateľovi pomocou PC s pripojením na sieť Internet, webovým prehliadačom a nainštalovaným LabVIEW Run-Time Engine prístup k implementovaným softvérovým prostriedkom na ovládanie optického vláknového refraktometrického systému cez Web. Hlavné okno tejto webovej stránky zobrazené po jej otvorení cez webový prehliadač je na Obr. 4.



Obr. 4 Webová stránka optického vláknového refraktometrického systému.

#### Skype

Na vizuálnu komunikáciu s meracím pracoviskom je využitý peer-to-peer program Skype, ktorý umožňuje prevádzkovať na to potrebné streamovanie videa. Sú tu k dispozícii aj internetová telefónia (VoIP), Instant messaging, ako aj možnosť prenosu súborov. Keďže je

voľne šíriteľný a poskytuje videotelefonovanie (mimo tradičných telefónnych sietí) medzi používateľmi zadarmo, je stále viac využívaný aj pre potreby vzdelávania. Pomocou nastaveného automatického odpovedania na prichádzajúce hovory na strane servera môžu vzdialení používatelia s programom Skype prijímať vizuálny obraz z meracieho pracoviska cez webovú kameru bez toho, aby musel byť niekto fyzicky prítomný na tomto pracovisku. Okno programu Skype použité pri vizuálnej komunikácii s meracím pracoviskom je zobrazené na Obr. 5.



Obr. 5 Okno programu Skype použité pri vizuálnej komunikácii s meracím pracoviskom.

### III. EXPERIMENTÁLNE OVERENIE SYSTÉMU

Pri experimentálnych meraniach boli použité dva senzory, a to senzor s LED diódou, pri používaní ktorého bol zvolený budiaci prúd 60 mA a ktorý ponorený v destilovanej vode sa považoval za čistý pri hodnote napätie na tomto senzore 9,1 V, a senzor s LD diódou, ktorý bol budený prúdom 10 mA a ktorý taktiež ponorený v destilovanej vode bol označený za čistý, ak detegoval napätie aspoň 4,1 V. Čo sa týka typov vykonaných experimentov, tieto boli trojaké. Prvý typ experimentov zahrňoval meranie indexu lomu známych a neznámych kvapalín pri teplote 20 °C. Pri druhom type bola meraná zmena veľkosti indexu lomu kvapaliny v rozsahu teplôt od 10 – 50 °C s krokom 5 °C. A tretí typ experimentálnych meraní sa sústredil na meranie zmeny indexu lomu kvapalín v závislosti od koncentrácie inej kvapaliny.

#### A. Meranie indexu lomu známych a neznámych kvapalín pri teplote 20 °C

Ako prvé boli vykonané merania indexu lomu známych kvapalín, ktorých výsledky spolu s ideálnymi hodnotami indexu lomu, ktoré boli získane v chemickom laboratóriu klasickými chromatografickými metódami, sú v Tab. 1 pre senzor s LED diódou a v Tab. 2 pre senzor s LD diódou. Tieto experimenty boli uskutočnené pri teplote meranej kvapaliny 20 °C. Chyby merania pri jednotlivých kvapalinách nepresahujú 0,5 %, z čoho je možné usudzovať, že optický vláknový refraktometrický systém funguje správne.

Tabuľka 1  
 Výsledky z merania indexu lomu známych kvapalín so senzorom s LED diódou

Meraná kvapalina	Nameraný index lomu	Ideálny index lomu	Chyba merania [%]
Destilovaná voda	1,3338	1,333	0,06
Technický lieh	1,3557	1,362	0,46
Stolový olej	1,4647	1,460	0,32

Tabuľka 2  
 Výsledky z merania indexu lomu známych kvapalín so senzorom s LD diódou

Meraná kvapalina	Nameraný index lomu	Ideálny index lomu	Chyba merania [%]
Destilovaná voda	1,3334	1,333	0,03
Technický lieh	1,3591	1,362	0,21
Stolový olej	1,4564	1,460	0,25

Pri experimentálnych meraniach neznámych kvapalín, t.j. kvapalín s neznámym indexom

lomu, boli použité tri kvapaliny, a to 50% pomarančový džús, 100% pomarančový džús a 100% jablkový džús. Pre tieto merania bol zvolený senzor s LED diódou. Namerané hodnoty indexu lomu týchto džúsov pri 20 °C sú uvedené v Tab. 3.

Tabuľka 3  
Výsledky z merania indexu lomu neznámych kvapalín so senzorom s LED diódou

Meraná kvapalina	Nameraný index lomu
50% pomarančový džús	1,3389
100% pomarančový džús	1,3419
100% jablkový džús	1,3399

#### B. Meranie indexu lomu kvapalín v závislosti od teploty

Ďalšie experimenty sa týkali merania zmeny indexu lomu pri zvyšovaní teploty skúmaných kvapalín. Úlohou týchto experimentov bolo zistiť, ako závisí hodnota indexu lomu predložených kvapalín od teploty. Pri týchto meraniach bol použitý senzor s LED diódou a skúmanými kvapalinami boli rôzne džúsy a technický lieh. Teplota bola zvyšovaná od 10 °C do 50 °C s krokom 5 °C. Získané výsledky z meraní ukazujú, že index lomu pre tieto skúmané kvapaliny pri zvyšovaní teploty lineárne klesá, a to rýchlosťou pre použité džúsy približne 0,00007/°C a pre technický lieh 0,0004/°C.

#### C. Meranie indexu lomu kvapalín v závislosti od koncentrácie

Posledným typom experimentov bolo meranie indexu lomu pri postupnej zmene koncentrácie jednej kvapaliny v inej kvapaline. Pre tieto experimentálne merania boli použité kvapaliny benzín (TEMPO PLUS 95) a destilovaná voda, ktoré boli samostatne miešané s technickým liehom s koncentráciou od 0 – 100 % a krokom 5 %. Teplota kvapalín bola udržiavaná na 20 °C. Z nameraných hodnôt vyplýva, že pre kombináciu kvapalín lieh – benzín platí lineárna zmena indexu lomu v závislosti od koncentrácie benzínu v technickom liehu so zmenou približne 0,0007/1%, a pre kombináciu voda – lieh je možné skonštatovať, že závislosť zmeny indexu lomu pri zvyšovaní koncentrácie liehu v destilovanej vode je nelineárna s maximálnou hodnotou indexu lomu pri 80% koncentrácii.

### IV. ZÁVER

Pretože na diaľku ovládané meracie pracoviská a s tým spojená možnosť efektívnejšieho vzdelávania sa tešia čoraz väčšiemu záujmu, je nepochybne veľkým prínosom aj nami predstavené unikátne zariadenie pre meranie indexu lomu kvapalín, t.j. laboratórny optický vláknový refraktometrický systém. Správnosť jeho fungovania bola experimentálnymi meraniami potvrdená s dosiahnutou chybou do 0,5 % pre LED senzor a do 0,25 % pre LD senzor.

### POĎAKOVANIE

Táto práca bola podporovaná Kultúrnou a edukačnou grantovou agentúrou (KEGA) Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky na základe projektu č. „063TUKE-4/2013 - Použitie diaľkovo ovládaného optického vláknového refraktometra vo výučbe“ a projektu č. „006TUKE-4/2014 - Experimentálny model PON siete TUKE a jeho použitie vo výučbe“. Táto práca je tiež výsledkom zmluvy „APVV-0025-12 - Predchádzanie vplyvu stochastických mechanizmov vo vysokorychlostných plne optických sieťach“ podporovanej Agentúrou na podporu výskumu a vývoja (APVV) Slovenskej republiky.

### LITERATÚRA

- [1] J. Bankoski, P. Wilkins, X Yaowu, „Technical Overview of VP8, an Open Source Video Codec for the Web“, *Multimedia and Expo (ICME)*, IEEE International Conference, 11-15 July 2011, doi: 10.1109/ICME.2011.6012227.
- [2] V. G. Gomes, et al., „Web-based Courseware in Teaching Laboratory-based Course“, *Global J. of Eng. Education*, vol.4, no.1, 2002, pp. 65-71.
- [3] J. Turán, L. Ovseník, *Projektovanie optických vláknových komunikačných a senzorových systémov*. Košice, Harlequin, 2007.
- [4] J. Turán, L. Ovseník, „Application of Web Controlled Fibre Optic Refractometer“, *IWSSIP 2014: 21st International Conference on Systems, Signals and Image Processing*, May 12-15, 2014, Dubrovnik, Croatia.

# Rozhranie k inteligentnému audiovizuálnemu archívu

<sup>1</sup>Tomáš KOCTÚR, <sup>2</sup>Matúš PLEVA, <sup>3</sup>Jozef JUHÁR

<sup>1, 2, 3</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>tomas.koctur@tuke.sk, <sup>2</sup>matus.pleva@tuke.sk, <sup>3</sup>jozef.juhar@tuke.sk

**Abstrakt** — Tento článok opisuje vybrané koncepty a riešenia použité pri vypracovaní rozhrania pre inteligentný audiovizuálny archív. Táto úloha je zameraná na vytvorenie rozhrania k inteligentnému archívu audiovizuálnych dát. Katedra elektroniky a multimediálnych telekomunikácií (KEMT) má funkčný serverový automatický prepis spojitkej reči do textu (ARR), ktorý po prijatí vstupného audio súboru vypíše časovo zarovnané rozpoznané slová. Zpracovanie užívateľského web rozhrania a automatické spracovanie týchto vstupov a výstupov do celkového systému bude opísané v tomto článku.

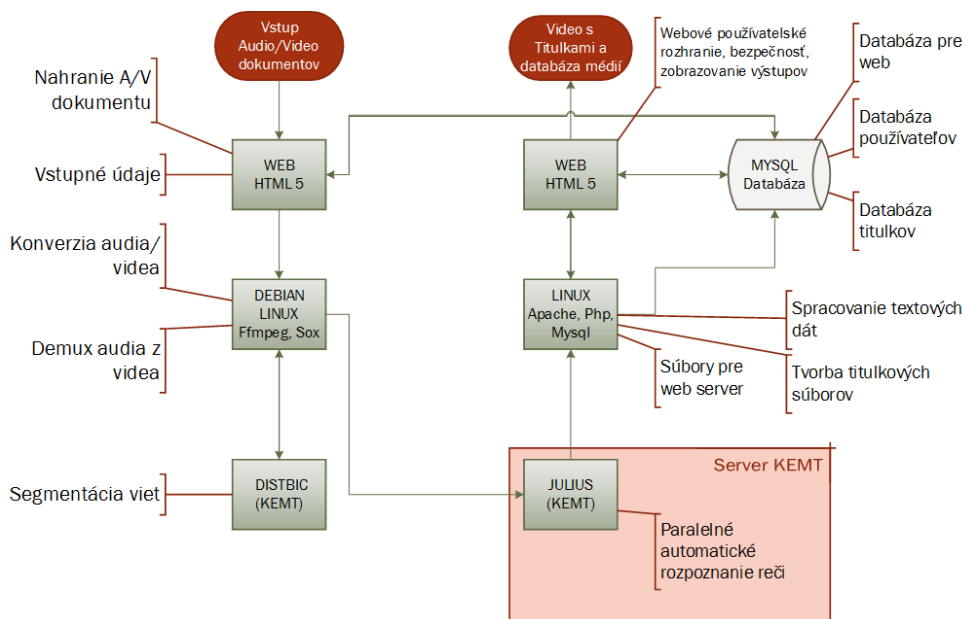
**Kľúčové slová** — automatické rozpoznávanie reči, MySQL, Unix Bash, titulkovanie

## I. ÚVOD

Systémy inteligentných audiovizuálnych archívov predstavujú aktuálny trend v oblasti multimédií [1] [2] [3] a stále zvyšujúceho sa množstva obsahu dostupného na webe.

Systém inteligentného audiovizuálneho archívu by mal prijímať vstupy vo forme audio alebo video dokumentov, ktoré sú zadávané cez webové rozhranie. Korešpondujúce výstupy ARR by mali byť na webe zobrazené vo videu vo forme titulkov (audio súbory sú konvertované videa pre účely titulkovania/anotácie). Takto vzniknutá databáza multimediálneho obsahu je ukladaná na server pričom odkazy na jednotlivé záznamy a príslušné titulky sú ukladané do databázy MySQL.

Ak by sme si jednoduchšie chceli predstaviť takýto systém (viď Obr.1) tak by išlo o webové rozhranie do ktorého sa dá vložiť (upload) video alebo audio dokument ktorý je týmto systémom spracovaný, textovo anotovaný. V databáze by bolo možné si takto textovo anotované video alebo audio prehrať s automaticky generovanými titulkami.



Obr.1 Bloková schéma systému.

## II. ÚVOD DO PROBLEMATIKY

Jadrom celého systému je server automatického rozpoznávania reči (ARR) [4] na Katedre elektroniky a multimediálnych telekomunikácii (KEMT), ktorý je fyzicky umiestnený v datacentre Ústavu výpočtovej techniky Technickej univerzity v Košiciach. Jadrom servera automatického rozpoznania reči je Julius [5] nainštalovaný na operačnom systéme Debian bez grafického rozhrania. Na pripojenie k tomuto rozpoznávaču je potrebný špecializovaný softvér na báze TCP: *SpeechRecoClient3*, ktorý zasiela audio vo formáte PCM wav so vzorkovacou frekvenciou 16kHz a s kvantovaním 16 bit na špecifickú službu na TCP/IP porte servera. Výstupom je rozpoznaná veta alebo súbor viet v textovej forme a časové zarovnanie jednotlivých slov (viď Obr. 2).

```
Connecting ... OK
Opening wav ... OK
Sending ... OK
Waiting response ... OK
<s> toto je vzorový výstup systému ARR pre účely tohto článku
</s> <s> trvalo to veľmi dlho </s>
=== begin forced alignment ===
-- word alignment --
id: from to n_score unit
-----
[ 0 1] -21.584606 [<s>] [<s>]
[ 2 162] -22.540953 [int] []
[ 163 216] -38.939487 [toto] [toto]
[ 217 250] -37.127167 [je] [je]
[ 251 266] -35.821228 [vzorový][vzorový]
[ 267 295] -36.879463 [systému][systému]
[ 296 322] -36.193867 [ARR] [ARR]
[ 323 401] -32.833736 [pre] [pre]
[ 402 431] -32.491894 [účely][účely]
[ 432 435] -32.117920 [tohto][tohto]
[ 436 457] -33.023617 [článku][článku]
[ 463 464] -46.309895 [</s>] [</s>]
=== end forced alignment ===
=== begin forced alignment ===
-- word alignment --
id: from to n_score unit
-----
[ 0 3] -2.584606 [<s>] [<s>]
[ 4 262] -22.540953 [trvalo] [trvalo]
[ 263 303] -35.939487 [to] [to]
[ 304 565] -38.127167 [veľmi] [veľmi]
[ 567 785] -35.821228 [dlho] [dlho]
[ 786 787] -56.309895 [</s>] [</s>]
=== end forced alignment ===
<EOR>
Exiting ...
```

Obr. 2 Vzor výstupu ARR systému.

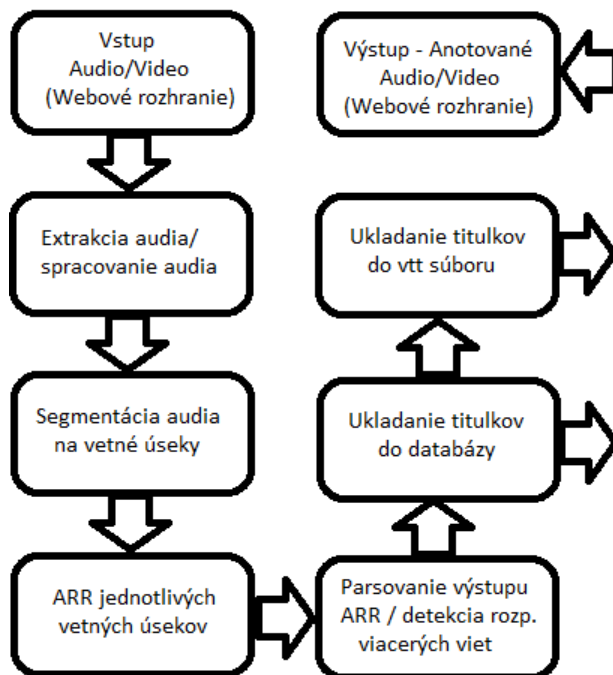
Z výstupu je možné vidieť, že pre jednotlivé slová majú svoj časový údaj začiatku a aj konca v stotinách sekundy. Z výstupu teda pomyselný prvý stĺpec reprezentuje začiatok, druhý koniec, tretí skóre rozpoznania, štvrtý typ výstupu alebo výstupné slovo a piaty výstupné slovo alebo znak. V štvrtom stĺpci sa môžu nachádzať znaky ako <s> alebo </s>, ktoré reprezentujú začiatok a koniec vety, značky ako *int* ktoré naznačujú výskyt šumu medzi slovami, alebo :@ reprezentujúce zaváhania alebo iné znaky označujúce typ výstupu ako napríklad <obec>.

Vstupné audio pre ARR by malo byť krátke s obmedzením na jednu vetu, čo bude mať za následok lepšie výsledky.

Na základe podmienok a odporúčaní pre vstup ARR a možnosťami výstupu je zjednodušená bloková schéma na Obrázku 1.

### A. Vstup audia a videa

Audio a Video sa vkladá cez webové rozhranie. Na túto časť nie sú potrebné žiadne špeciálne nástroje a postačuje obyčajná PHP web-stránka. Avšak ak by bola požiadavka zobrazovať údaje o stave uploadu, tak je vhodné použiť rozšírenie pre PHP server APC, ktorý rozširuje RFC-1867 o uploadu z polí [6]. V tomto úvodnom kroku je potrebné definovať aj vstupné názvy médií, vstupné premenné a iné, pričom je dobré ich zapisovať priamo do databázy.



Obr.3 Zjednodušená bloková schéma systému.

### B. Extrakcia audia a jeho spracovanie

Vstupom systému môže byť audio alebo video. Z videa je potrebné extrahovať/demultiplexovať iba audio stopu. V oboch prípadoch však treba naplniť vstupné špecifikácie pre systém ARR čo je vzorkovacia frekvencia  $16\text{ kHz}$  a  $16\text{ bit}$  kvantovanie. K tomu všetkému má byť takéto extrahovanie automatické a tak je vhodné použitie konzolovej aplikácie na báze OS Linuxu bez grafického rozhrania. Najvhodnejším open source riešením je OS *Debian*. Nástroje pre tento OS, ktoré boli použité sú: jeho *shell*, *ffmpeg* a *sox*. Druhý menovaný má širokú podporu komunity a preto sú podporované aj moderné formáty videa čo bolo jednou z hlavných požiadaviek. Nástroj *sox* je používaný na automatickú normalizáciu audia.

Keďže neexistuje celoplošné jednotné riešenie na zlepšenie kvality zvuku, tak je normalizácia v tomto systéme len doplnková na rôzne testovacie účely. Vhodnejšie by bolo použitie dvojnásobného spektrálneho odčítania šumu, čo by malo zlepšiť výsledky oproti jednoduchej normalizácii.

### C. Parsovanie výstupu ARR a detekcia rozpoznania viacerých viet

Kľúčovú úlohu hrá segmentácia audia. Keďže systém nemá mať obmedzenia na dĺžku vstupného audia alebo videa, samozrejme v reálnych diskových obmedzeniach, tak by boli takéto dlhšie audio-video rozpoznávané veľmi dlho, resp. by neboli rozpoznané vôbec. ARR systém totiž disponuje timeout periódou (ktorú spôsobí prídlhé spracovanie dlhého úseku, resp. spadnutie rozpoznávania, kvôli prekročeniu limity na dĺžku jednej hypotézy), takže nie je možné rozpoznávať úseky dlhšie ako 10minút.

Čím lepšie segmentovaný vetný úsek je poslaný na rozpoznanie tým kvalitnejšie je rozpoznanie, keďže ARR systém aplikuje aj jazykový model slovenského jazyka (vetných konštrukcií). V kratších vetách nemusí jazykový model spôsobiť zlepšenie rozpoznania.

Na rozdelenie na kratšie úseky je použitý softvér *distbic* (vlastná implementácia DISTBIC algoritmu [7]), ktorý rozpoznáva zmenu hovoriaceho v audio alebo deteguje dlhšie medzivetné pauzy. Výstupom tohto softvéru je označenie časové začiatku a konca segmentov.

0.00 4.50 speech 1  
4.50 18.90 speech 2  
18.90 21.90 speech 3  
21.90 26.40 speech 4  
26.40 32.60 speech 5  
32.60 35.30 speech 6  
35.30 40.20 speech 7  
40.20 43.50 speech 8  
43.50 45.40 speech 9  
45.40 62.18 speech 10

Obr. 4. Vzor výstupu distbicu

Časové zarovnanie je v sekundách s rozlíšením na stotiny sekundy.

Rozdelenie audia na kratšie audio segmenty je realizované pomocou vyššie spomínaného nástroja *ffmpeg*, ktorý ma oproti nástroju *sox* výhodu využívania viacerých jadier procesora naraz čo má za následok rýchlejšie spracovanie audia na servery s viacerými jadrami.

Na rozdelenie v nástroji *ffmpeg* je potrebné konvertovať čas v sekundovom formáte na formát hh:mm:ss.milisekundy. Na takúto konverziu sa je použitý *bash* script v prostredí OS *Debian*. Netreba zabudnúť na vhodné názvoslovie výstupných súborov, kde je vhodné v pomenovaní zakomponovať poradie jednotlivých audio segmentov.

#### D. ARR v jednotlivých vetných úsekoch

Audio vetné úseky sa vkladajú do ARR systému, nainštalovanom na dedikovanom servery KEMTu, cez prístupový softvér SpeechRecoClient3 ARR server podporuje paralelné spracovanie požiadaviek preto je možné zasielať viaceré požiadavky naraz. Sever v súčasnosti podporuje spracovanie až 12 požiadaviek naraz. Požiadavky na rozpoznanie audio súborov rôznych dĺžok sú spracované počas rôznych časových intervalov. Ak je požiadavka na to aby boli zasielané na ARR server napríklad 4 naraz je potrebné automaticky sledovať či ich rozpoznanie už bolo dokončené. SpeechRecoClient3 sa po rozpoznaní jednotlivého audio súboru ukončí. Systém inteligentného archívu teda sleduje jednotlivé spustené inštancie SpeechRecoClient-a a zapisuje si ich PID (Process identifier). Pomocou zobrazenia príkazu „ps aux“ je možné sledovať jednotlivé spustené procesy a ich PID. Jednoduchým bash skriptom je možné monitorovať či procesy jednotlivých inštancií SpeechRecoClient-a sú už ukončené alebo ešte pracujú. Výhodou 3 verzie tohto klienta je timeout perióda, kde je rozpoznanie po 10 minútach ukončené bez výsledku.

Výstupom ARR je text ako na Obr. 2. Takýto výstup je potrebné automaticky parsovať a vyberať z neho jednotlivé slová a k nim korešpondujúce časové zarovnania. Výstupom môže byť aj séria segmentov a v segmentoch môžu byť označené miesta kde je potencionálny začiatok novej vety. Spracovanie parsovaním je opísané v samostatnej Kapitole III.

#### E. Ukladanie titulkov do databázy

Výstupom systému je video aj v prípade že vstupom bolo audio. Audio je tým pádom konvertované na video so statickým obrázkom. Toto riešenie bolo implementované kvôli tomu, že žiadne známe audio web prehrávače nepodporujú titulky.

Po videom sú zobrazované titulky ktoré sú uložené v databáze. HTML5 player vo webovom rozhraní podporuje zobrazovanie aktuálneho času prehrávaného videa, vďaka čomu sa mohla vytvoriť história zobrazených tituliek a ich následné upravenie ale spracovanie. Taktiež je pridaná funkcia toho, že po kliknutí na titulok v histórii je video nastavené na začiatkový čas titulku. Nevýhodou zobrazenia titulkov pod videom je to že pri zobrazení na celú obrazovku nie je možné používať tieto funkcie. Výhodou je to, že v prípade použitia mobilných zariadení ako sú smartphony alebo tablety či PDA, kde nie sú natívne podporované zobrazenia tituliek z externého súboru je takéto riešenie jedinou možnosťou. Hlavou funkciou inteligentného archívu audiovizuálnych dát je vyhľadávanie kľúčových slov v titulkoch. Bez uložených titulkov v databáze by to bolo náročnejšie. Na to aby tieto funkcie bolo možné vytvoriť je potrebné jednotlivé vetné segmenty zapisovať do databázy s ich začiatkovými a koncovými časom.

Ako databázový systém bola použitá MySQL databáza, ktorá je vhodným open source riešením s jednoduchou implementáciou a postačujúcimi funkciami a indexovacími metódami.



#### F. Ukladanie titulkov do vtt súboru

Nový titulkový súborový formát *vtt* (web video text track), má nahrádzať doterajší *srt* formát, kde jeho hlavnou úlohou je štandardizovať zápis titulkov do súboru a aj pridáva nové „exotické“ funkcie ako je karaoke zobrazenie a taktiež ruší niektoré limity predošlého formátu. Vo formáte *vtt* už nie je potrebné očíslovať jednotlivé titulky, tie sú načítavané aj ak nie sú v správnom chronologickom poradí. Pridané boli funkcie zobrazenia farby a štýlu textu. Pridané boli možnosti zarovnania a pozície textu ako aj zobrazenia sprava doľava, príznačné pre arabské krajiny, a aj zobrazenia z vrchu dole, príznačné pre ázijské krajiny.

WEBVTT

00:00:01.000 --> 00:00:10.000

This is the first line of text, displaying from 1-10 seconds

00:00:15.000 --> 00:00:20.000

And the second line of text  
separated over two lines

Obr. 5. Vzor syntaxu vtt súboru.

WEBVTT

00:00:01.000 --> 00:00:10.000

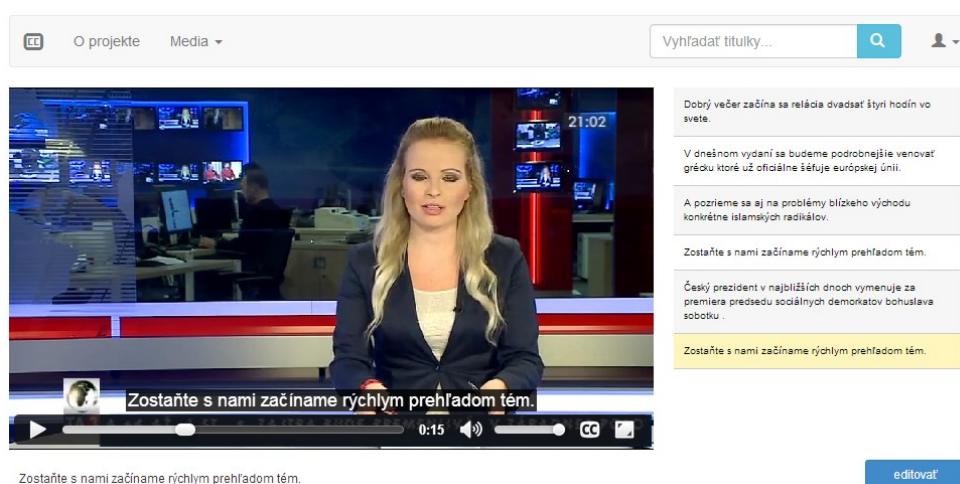
Never gonna give you up <00:00:01.000> Never gonna let you down  
<00:00:05.000> Never gonna run around and desert you

Obr. 6. Vzor syntaxu vtt súboru pre karaoke zobrazenie.

Výstupom parsera sú jednotlivé súbory v ktorých je už časová informácia a text. Tieto súbory bolo potrebné jednoducho spojiť, čo bolo realizované pomocou bash scriptu. Viac v kapitole III.

#### G. Prezentovanie anotovaného videa vo web rozhraní

Na Obrázku 7 môžeme vidieť ukážku systému. Vo webovom rozhraní je skonvertované video do formátu MP4 zobrazované s titulkami vo formáte *vtt*. Ako video player je použitý základný moderný HTML5 player. Na zobrazenie nie je potrebné inštalovať žiadne doplnky, stačí mať aktualizovaný moderný web prehliadač.



Obr. 7. Ukážka systému.

Pod videom sú zobrazené aktuálne titulky uložené v databáze zobrazené podľa aktuálneho timestampu videa. V pravej časti sa nachádza história titulkov. V histórii titulkov je možné kliknúť na titulok a video sa nastaví na korešpondujúci čas. Titulky v databáze je možné editovať pričom výstupy po editácii je možné stiahnuť do súboru pre účely zdokonaľovania ARR systému, či pre použitie s originálnym videom.



Vo videu sú zobrazované titulky z titulkového súboru *vtt*, ktorých limitácia je, že súbor je načítaný iba pri načítaní stránky a po potencionálnej editácii tituliek v databáze je potrebné znovu načítať stránku (vygeneruje sa aktualizovaný titulkový súbor).

### III. PARSOVANIE VÝSTUPU ARR A DETEKCIA ROZPOZNANÝCH VIET

Výstup ARR je zobrazený na Obrázku 1. Symbol `<s>` a `</s>` označujú začiatky a konce viet. Veta je väčšinou úsek audio súboru poslaného na rozpoznanie, teda nie skutočná veta. Štandardne je jedna veta na celý audio súbor ale v špecifických prípadoch ich môže byť viac. V prípade rozpoznania viacerých viet má každá veta v rámci jedného výstupu vlastné časovanie začínajúce od nuly. V rámci jedného segmentu môže byť detegovaná pauza, čo môžeme označiť ako novú vetu, kvôli lepšej čitateľnosti tituliek. Pauzu môže vyjadrovať napríklad značka `[int]` (intermittent noise).

Parser teda z výstupu rozpoznávača najprv postupne načíta počet viet a potom počet menších vetných segmentov v rámci jedného audio segmentu. Rozpoznaný audio súbor ma nejaký začiatkový čas, ktorý je výstupom *distbic*-u (viď. Obr. 2). Tento čas je potrebné prirátat k časom v segmente. V prípade, že vo výstupe je rozpoznaných viac viet tak je potrebné prirátat, už prirátaný čas, posledného `</s>` element segmentu, k počiatkovému `<s>` elementu nasledujúceho segmentu a následne to prirátat ďalej k jednotlivým položkám v rámci segmentu. Tento postup je potrebné aplikovať na každý rozpoznaný segment, aby sa zabezpečila správnosť časových údajov.

Parser následne načítava každý riadok zvlášť a pripisuje jednotlivé slová do stringu. Z načítavania sú vyradené špeciálne symboly ako je napríklad `:@`. Keď je načítaná celá veta, tak je upravený prvý znak vety na veľké písmeno. Z prvého a posledného riadku vety je vybraný upravený časový údaj a je z neho spravená hlavička *vtt* súboru. Parser označuje názvoslovie jednotlivých súborov poradovým číslom aby potom bash skript na spojenie jednotlivých rozpoznaní jednoducho cyklicky spojil všetky parsované súbory.

### IV. ZÁVER

V tomto článku sú v skratke opísané základy systému inteligentného multimedialného archívu. Systém je dostupný na webovej stránke: <http://sub.radio9.sk/>.

V súčasnosti je systém funkčný a pracuje bez väčších problémov. Kvôli obmedzenej diskovej kapacite je prístup do systému obmedzený. Systém je monitorovaný a v prípade zlyhania je informovaný správca. Takéto incidenty sú priebežne opravované a vďaka nim je systém čoraz viac stabilný.

### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Kompetenčné centrum znalostných technológií pre inovácie produkčných systémov v priemysle a službách (kód ITMS projektu: 26220220155) 50% a projektu Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií (ITMS-26220220182) 50% na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

### REFERENCIE

- [1] V. Cabarrão et al., „Revising the Annotation of a Broadcast News Corpus: a Linguistic Approach“, 9<sup>th</sup> International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, 2014, pp. 3908 – 3913.
- [2] M. Elmahdy, M. Hasegawa-Johnson, E. Mustafawi, „Development of a TV Broadcasts Speech Recognition System for Qatari Arabic“, 9<sup>th</sup> International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, 2014, pp. 3057 – 3061.
- [3] A. Zgank, A. Z. Vitez, D. Verdonik, „The Slovene BNSI Broadcast News Database and Reference Speech Corpus GOS: Towards the Uniform Guidelines for Future Work“, 9<sup>th</sup> International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, 2014, pp. 2644-2647.
- [4] S. Ondáš, M. Pleva, M. Lojka, M. Sulír, J. Juhár, „Server-based speech technologies for mobile robotic applications“, Journal of Electrical and Electronics Engineering. Vol. 6, no. 1 (2013), p. 95-98. ISSN 1844-6035.
- [5] A. Lee, T. Kawahara, „Recent Development of Open-Source Speech Recognition Engine Julius“, Proceedings of the Asia-Pacific Signal and Information Processing Association, Annual Summit and Conference, APSIPA ASC 2009, Sapporo, Japan, pp. 131-137.
- [6] RFC 1867, *Form-based File Upload in HTML*, [www.ietf.org/rfc/rfc1867.txt](http://www.ietf.org/rfc/rfc1867.txt).
- [7] A. Vandecatseye, J.-P. Martens. "A fast, accurate and stream-based speaker segmentation and clustering algorithm" Proceedings of INTERSPEECH, 2003, pp. 941-944.

# Algoritmus rozdelenia kódovej knihy v procese vektorového kvantovania s trelisovým kódovaním

Iveta GLADIŠOVÁ

Katedra elektroniky a multimediálnych telekomunikácií  
Fakulta elektrotechniky a informatiky, Technická univerzita Košice, SR

iveta.gladisova@tuke.sk

**Abstrakt** — Článok sa zaoberá algoritmom rozdelenia súboru kvantizačných vektorov, t.j. kódovej knihy, v procese vektorového kvantovania s trelisovým kódovaním. Najprv je opísaný princíp vektorového kvantovania s trelisovým kódovaním a následne je rozobraný algoritmus rozdelenia jeho kódovej knihy. V programovom jazyku C++ bol zrealizovaný simulačný program vektorového kvantizátora s trelisovým kódovaním s uvedeným algoritmom rozdelenia kódovej knihy. V článku sú predstavené výsledky vplyvu použitia rôznych štruktúr treliasa na hodnoty celkového pomeru signálu k šumu tohto kvantizátora.

**Kľúčové slová** — vektorové kvantovanie, trelisové kódovanie, rozdelenie kódovej knihy.

## I. ÚVOD

Základnou vlastnosťou vektorového kvantovania s trelisovým kódovaním je použitie štruktúrovanej kódovej knihy s rozšíreným súborom kvantizačných vektorov [3], [1] a [6]. Štruktúra treliasa potom rozšírený počet kvantizačných vektorov obmedzí podľa požadovanej prenosovej rýchlosti. Princíp spočíva v tom, že pri klasickom vektorovom kvantovaní sa ku vstupnému vektoru musí nájsť taký kvantizačný vektor, ktorý minimalizuje jeho výber s čo najmenším možným skreslením na aktuálnom časovom intervale. To niekedy môže viesť do stavu nevhodného pre nasledujúci vstupný vektor a spôsobiť tak väčšie výsledné skreslenie v dlhšom časovom intervale. Namiesto porovnávania jedného vstupného vektora so súborom kvantizačných vektorov obsiahnutých v kódovej knihe a produkovania výsledného kanálového slova, môže kodér súčasne uvažovať napr. dva nasledujúce vstupné vektory naraz, nájsť najmenšie výsledné skreslenie pre dva výstupné (kvantizačné) vektory a vyslať ich prislúchajúce kanálové slová. To zaisť kodéru menšie minimálne skreslenie pre dva za sebou idúce vektory, než pri ich jednotlivom kódovaní [8]. Všeobecne môže kodér uvažovať  $L$  za sebou idúcich vstupných vektorov. Potom nájde spomedzi všetkých možných postupností kanálových symbolov takú, ktorej zodpovedajúca postupnosť vektorov z kódovej knihy dáva najmenšie celkové skreslenie medzi  $L$  vstupnými a  $L$  výstupnými symbolmi [2], [5]. Článok sa zameriava na algoritmus rozdelenia kódovej knihy pri vektorovom kvantovaní s trelisovým kódovaním a predstavuje dosiahnuté výsledky pomocou zrealizovaného simulačného programu vektorového kvantovania s trelisovým kódovaním pre konkrétne štruktúry treliasa s cieľom zistenia jeho účinnosti.

## II. VEKTOROVÉ KVANTOVANIE S TRELISOVÝM KÓDOVANÍM

Vektorový kvantizátor s trelisovým kódovaním (VKTK) je automat s konečným počtom stavov  $s_i$ , so stavovým priestorom  $S = \{s_i : i = 0, 1, 2, \dots, 2^{v-1}\}$ , súborom výstupných symbolov  $Y = \{0, 1\}^{vn}$  (t.j. množina všetkých binárnych slov dĺžky  $vn$ ), kde  $v$  je rozmer euklidovského priestoru  $R_v$  a  $n$  je počet bitov na zložku vektora, ďalej s kódovou knihou  $B = \{\bar{b}_i \in R_v : i = 0, 1, 2, \dots, S^{v-1} - 1\}$ , kde  $\bar{b}_i$  sú kvantizačné vektory, s funkciou nasledujúceho stavu a výstupnou funkciou  $d$  [4]. Funkcia nasledujúceho stavu a výstupu, teda tzv. trelis, pri VKTK dekodéri má špeciálny formát. Predpokladajme, že dekodér je v danom stave v čase  $t$ . Potom kodér môže vyslať ľubovoľné z  $2^{v+n'}$  možných binárnych slov dĺžky  $vn$ , ktoré sú rozdelené do  $D_i$  podsúborov [2], [9]. Potom prvé  $n'$  bity vysielaného slova určujú stav

dekodéra v čase  $t+1$  a spôsobia, že dekodér prejde do tohto stavu. Zvyšné  $vn - n'$  bity určujú výstup dekodéra v čase  $t$ . To spôsobí, že forma zmeny z jedného určitého stavu do iného je určená ľubovoľným z  $2^{vn-n'}$  binárnych slov dĺžky  $vn$ . Táto zmena je reprezentovaná v treliisovom diagrame jednou vetvou, namiesto  $2^{vn-n'}$  paralelných zmien. Preto je  $2^{n'}$  vychádzajúcich vetiev z každého stavu (uzla). S touto konverziou môže každá vetva treliisa zo subknihy (podsubor kódovej knihy  $B$ ) byť označená binárnym slovom dĺžky  $n'$ . Každá subknih  $D_i$  je súbor kvantizačných vektorov zapísaných mimo dekodéra pomocou posledných  $vn - n'$  bitov z celého binárneho slova vysielaného kódom. Pretože každá táto subknih je veľkosti  $2^{vn-n'}$ , musí byť v treliise práve  $2^{n'+M}$  subkníh  $D_i$ , aby bola do nich rozdelená celá kódová kniha  $B$  ( $M$  je celočíselná konštanta, uvažovalo sa  $M=1$ ).

Funkciou VKTK kodéra je nájsť najlepšiu možnú reprodukciu pre ľubovoľnú danú postupnosť zdrojových vektorov (dĺžky  $L$ ). Presnejšie, nájsť zodpovedajúcu postupnosť kódových vektorov v  $S$ -stavovom treliise s kritériom minimálnej euklidovskej vzdialenosti. Každá postupnosť kódovaných vektorov v treliise je popísaná výslednou cestou, tj. treliisom a výberom kódovaných vektorov z kódovej knihy priradených ku vetvám treliisa. To vyžaduje  $Ln'$  bitov pre prvú časť (určenie stavu) a  $L(vn - n')$  bitov pre časť druhú (určenie vetvy treliisa), ktorá sa pridáva k celkovému počtu  $Lvn$  bitov [4]. Pre danú postupnosť zdrojových vektorov kodér používa Viterbiho algoritmus na hľadanie najlepšej postupnosti kódovaných vektorov v treliise [1], [7]. Potom nasleduje tzv. vypúšťanie všetkých  $Lvn$  bitov určujúcich túto postupnosť na konci  $S$ -stavového treliisa, ktorý používa dekodér na reprodukovanie postupností. Tento proces je opakovaný pre ďalšie nasledujúce  $L$  zdrojové vektory.

### III. VŠEOBECNÝ ALGORITMUS DELENIA KÓDOVEJ KNIHY PRE VKTK.

Cieľom delenia kódovej knihy  $B$  používanej vo VKTK je získať podsubory  $D_i$  kvantizačných vektorov s maximalizovanou minimálnou euklidovskou vzdialenosťou v rámci každého podsuboru (táto vzdialenosť medzi  $\bar{b}_i$  je na obr. 1 znázornená rovnakou farbou vektorov). Podsubory sú potom priradené vetvám treliisa podľa predpisu uvedeného v [2].

Uvažujme kódovú knihu  $B$  veľkosti  $N = 2^{vn+n'}$  kvantizačných vektorov  $\bar{b}_i$ , kde  $v$  je rozmer vektora a  $n$  je počet bitov na vzorku. Najprv sa vypočítajú euklidovské vzdialenosti medzi všetkými možnými párami kvantizačných vektorov a zoradia sa do neklesajúcej postupnosti spolu so zodpovedajúcimi vektormi. To dáva tabuľku s  $N \times (N-1)/2$  údajmi, kde  $i$ -tému vstupu zodpovedá pár kódovaných vektorov  $\bar{b}_i$  a  $\bar{b}'_i$ , ktorých vzdialenosť je  $d_i = \|\bar{b}_i - \bar{b}'_i\|_2$ . Prvé  $\bar{b}_0$  a  $\bar{b}'_0$  sú umiestnené v tzv. „global“ súboroch  $A_0$  a  $A_1$  jednotlivo, čím je prvý vstupný pár kvantizačných vektorov vybraný z tabuľky. Potom tento základný krok je opakovaný s ďalšími zoradenými párami vektorov dovtedy, kým veľkosť jedného zo súborov  $A_0$  alebo  $A_1$  nedosahuje hodnotu  $N/2$ . Všetky dvojice vektorov, ktoré už boli priradené, sú tak vybrané zo zoradzovacej tabuľky. Tie vektory, ktoré ešte neboli zadelené, sa zadedia do druhého „global“ súboru.

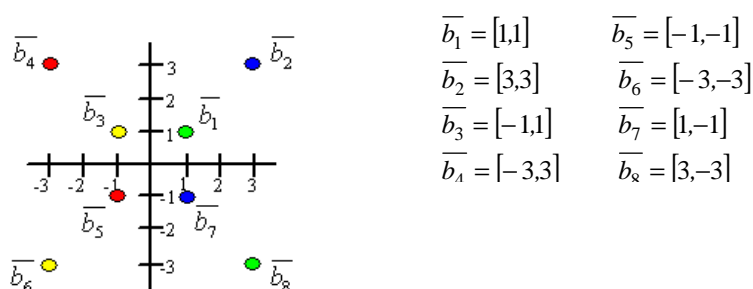
**Základný algoritmus** delenia kódovej knihy je formálne popísaný nasledovne [2], [4]:

- 1) Vyhľadaj v tabuľke index  $j$  taký, že platí  $\forall i < j$  ani  $\bar{b}_i$ , ani  $\bar{b}'_i$  nepatria do  $A_0 \cup A_1$ , ale aspoň jeden z  $\bar{b}_j$  a  $\bar{b}'_j$  patrí do  $A_0 \cup A_1$ .
- 2) Ak obidva  $\bar{b}_j$  a  $\bar{b}'_j$  patria do  $A_0 \cup A_1$ , t.j. už sú obidva zadelené, potom  $\bar{b}_i \leftarrow \bar{b}_{i+1}$ ,  $\bar{b}'_i \leftarrow \bar{b}'_{i+1}$  a  $d_i \leftarrow d'_{i+1}$  pre  $\forall i \geq j$ . Pokračuj krokom 1).
- 3) Ak  $\bar{b}_j$  patrí do  $A_0$  (alebo  $A_1$ ), zadel'  $\bar{b}'_j$  do  $A_1$  (alebo  $A_0$ ). Ak  $\bar{b}'_j$  patrí do  $A_0$  (alebo  $A_1$ ), zadel'  $\bar{b}_j$  do  $A_1$  (alebo  $A_0$ ).
- 4) Ak veľkosť  $A_0$  (alebo  $A_1$ ) dosiahne  $N/2$ , potom pridaj zvyšné nezadelené kódové vektory (ak sú také) do  $A_1$  (alebo  $A_0$ ) a ukonči proces. Ináč pokračuj krokom 1).

Na rozdelenie celej kódovej knihy  $B$  veľkosti  $N$  do  $2^{n'+M}$  podsuborov  $D_i$  sa aplikuje opísaný algoritmus, čím sa získajú najprv 2 globalsúbory  $A_0$  a  $A_1$  veľkosti  $2^{vn+n'-1}$ . Tieto dva súbory sa ďalej rozdelia tým istým algoritmom. Výsledkom budú 4 podsubory  $D_i$  veľkosti  $2^{vn+n'-2}$ . To sa opakuje, kým ich počet nie je  $2^{n'+M}$ , čím je rozdelenie celej kódovej knihy ukončené.

Výpočtová náročnosť algoritmu delenia kódovej knihy kvantizačných vektorov do podsúborov  $D_i$  je nasledovná: A) pri výpočte euklidovských vzdialeností sa vypočítajú metriky medzi všetkými vektormi tzn. vyčíslí sa výraz:  $\sum_{k=1}^v (x_{ik} - x_{jk})^2$ ,  $i, j = 1, 2, \dots, N$ , kde  $v$  je rozmer vektora,  $i, j$  sú indexy vektorov a  $N$  je počet vektorov kódovej knihy. V tomto kroku sa vykoná  $v \cdot N \cdot (N-1)/2$  odčítaní a násobení. B) Pri zoradzovaní vypočítaných metrik sa v každom kroku zoradzovania musí prejsť celý súbor a vybrať najmenší prvok z  $N \cdot (N-1)/2$  prvkov. Tento krok prebieha dovtedy, pokiaľ sa nevyskytnú indexy všetkých vektorov kódovej knihy. Maximálny počet operácií môže byť  $\frac{N \cdot (N-1)}{2} \cdot \frac{N \cdot (N-1)}{2} - 1$  porovnaní. C) Operácia priradzovania vektorov do podsúborov  $D_i$  si vyžaduje v každom cykle  $N/2$  porovnaní a jedno priradenie.

Uvažujme napr. nasledujúcu 2-rozmernú ( $v=2$ ) celočíselnú kódovú s  $N = 8$  kvantizačnými vektormi, na ktorej by sme ukázali algoritmus jej rozdelenia do prislúchajúcich podsúborov  $D_i$  s následným zakódovaním podľa VKTK.



Obr. 1 Príklad kódovej knihy  $B$  veľkosti  $N = 8$  kvantizačných vektorov  $\bar{b}_i$ .

1) Určenie euklidovských vzdialeností  $v$ -rozmerných vektorov všetkých možných párov  $\bar{b}_i$  a  $\bar{b}_i'$ :

$$\overline{b_1 b_2} = \sqrt{(1-3)^2 + (1-3)^2} = 2\sqrt{2} = 2,83$$

$$\overline{b_1 b_3} = \sqrt{(1+1)^2 + (1-1)^2} = \sqrt{4} = 2$$

$$\overline{b_1 b_4} = \sqrt{(1+3)^2 + (1-3)^2} = \sqrt{20} = 4,47$$

....atď, až

$$\overline{b_7 b_8} = \sqrt{(1-3)^2 + (-1+3)^2} = \sqrt{8} = 2,83$$

2) Zoradenie vzdialeností vektorov do neklesajúcej postupnosti:

$$d_0 : \overline{b_1 b_3} = 2$$

$$d_1 : \overline{b_1 b_7} = 2$$

$$d_2 : \overline{b_3 b_5} = 2$$

....atď, až

$$d_{27} : \overline{b_4 b_8} = 8,49$$

3) Delenie kódovej knihy  $B$  veľkosti  $N$  do  $2^{n'+M}$  podsúborov  $D_i$ , pre náš prípad s  $N = 2^{vn+n'} = 2^{2 \cdot 1+1} = 8$  kvantizačnými vektormi znázornených na obr.1 sa vykoná do  $2^{n'+M}$  t.j.  $2^{1+1} = 4$  podsúborov  $D_i$ , t.j.  $(D_0, D_1, D_2, D_3)$  a to nasledovným spôsobom:

- V prvom kroku sa kódová kniha  $B$  rozdelí do dvoch globalsúborov  $A_0$  a  $A_1$ , každý o veľkosti  $2^{vn+n'-1} = 2^{2 \cdot 1+1-1} = 4$  kvantizačných vektorov. Toto delenie sa opakuje, kým nie sú všetky kvantizačné vektory zadené do  $2^{n'+M} = 4$  podsúborov  $D_i$ .
- V druhom kroku sa globalsúbory  $A_0$  a  $A_1$  rozdelia už popísaným postupom v 1) a 2), čím dostaneme požadovaný počet 4 podsúborov  $D_i$ , každý o veľkosti  $2^{vn+n'-2} = 2^{2 \cdot 1+1-2} = 2$  kvantizačné vektory. Uvedené delenie sa teda opakuje dovtedy, kým počet podsúborov nie je  $2^{n'+M}$ . Nakoľko pre náš prípad sme už zadelili vektory do prislúchajúceho počtu podsúborov (na základe vzťahu  $2^{n'+M} = 2^{1+1} = 4$  podsúbory  $D_i$ ), potom proces delenia kódovej knihy  $B$  je ukončený.

$$A_0 = \{\bar{b}_1, \bar{b}_5, \bar{b}_4, \bar{b}_8\} \Rightarrow \begin{aligned} d_0 : \bar{b}_1 \bar{b}_5 &= 2.83 \\ d_1 : \bar{b}_1 \bar{b}_4 &= 4.47 \\ &\vdots \\ d_5 : \bar{b}_4 \bar{b}_8 &= 8.49 \end{aligned} \Rightarrow \begin{aligned} D_0 &= \{\bar{b}_1, \bar{b}_8\} \\ D_2 &= \{\bar{b}_5, \bar{b}_4\} \end{aligned}$$

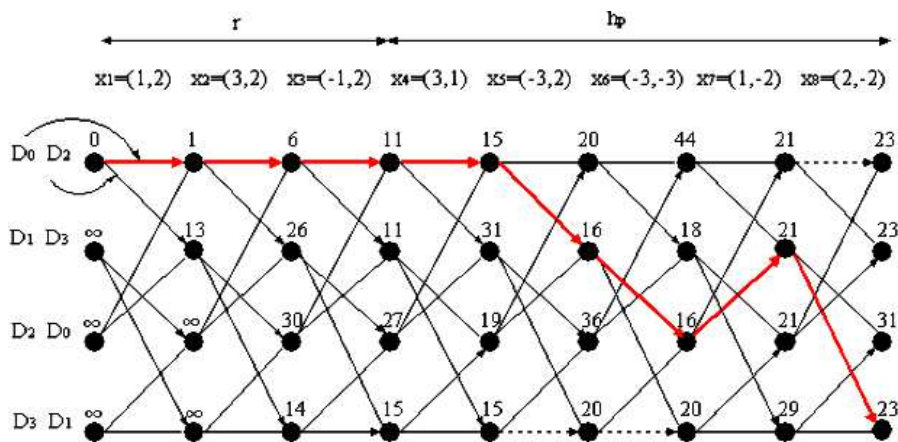
$$A_1 = \{\bar{b}_3, \bar{b}_7, \bar{b}_2, \bar{b}_6\} \Rightarrow \begin{aligned} d_0 : \bar{b}_3 \bar{b}_7 &= 2.83 \\ d_1 : \bar{b}_2 \bar{b}_3 &= 4.47 \\ &\vdots \\ d_5 : \bar{b}_2 \bar{b}_6 &= 8.49 \end{aligned} \Rightarrow \begin{aligned} D_1 &= \{\bar{b}_3, \bar{b}_6\} \\ D_3 &= \{\bar{b}_7, \bar{b}_2\} \end{aligned}$$

Po uvedenom delení vektorov do jednotlivých podsúborov tak dostaneme 4 podsúbory ( $D_0, D_1, D_2, D_3$ ), každý veľkosti 2 kvantizačné vektory, konkrétne  $D_0 = \{\bar{b}_1, \bar{b}_8\}$ ,  $D_2 = \{\bar{b}_4, \bar{b}_5\}$ ,  $D_1 = \{\bar{b}_3, \bar{b}_6\}$ ,  $D_3 = \{\bar{b}_2, \bar{b}_7\}$ , pričom vzdialenosti dvoch kvantizačných vektorov s maximalizovanou minimálnou vzdialenosťou prislúchajúcich jednotlivým podsúborom  $D_i$  sú na obr.1 znázornené farebne. Tieto podsúbory sú potom priradené jednotlivým vetvám treliisa [4].

Príklad vektorového kvantovania s treliisovým kódovaním pre 4-stavový treliis konkrétne pre štruktúru s  $M=1$  (t. j. dvojvetvové) a pre uvedenú kódovú knihu s  $2^{2+1} = 8$  kvantizačnými vektormi (obr.1) je znázornený na obr.2. Parametre kodéra boli nasledujúce: hĺbka prehľadávania  $h_p = 5$  a oneskorenie  $r = 3$ . Parametre treliisa boli nasledovné:  $v = 2$ ,  $n = 1 \text{ bit / vzorku}$ ,  $n' = 1$  a  $M = 1$ .

Ďalej uvažujme nasledujúcu postupnosť náhodných vstupných vektorov  $x_1 = (1, 2), x_2 = (3, 2), x_3 = (-1, 2), x_4 = (3, 1), x_5 = (-3, 2), x_6 = (-3, -3), x_7 = (1, -2), x_8 = (2, -2)$

a podsúbory  $D_0 = \{\bar{b}_1, \bar{b}_8\}$ ,  $D_2 = \{\bar{b}_4, \bar{b}_5\}$ ,  $D_1 = \{\bar{b}_3, \bar{b}_6\}$ ,  $D_3 = \{\bar{b}_2, \bar{b}_7\}$ . Pre uvedené parametre a štruktúru treliisa sa potom VKTK vykoná podľa uvedeného postupu.



Obr.2. Príklad vektorového kvantovania s treliisovým kódovaním s vyznačenou optimálnou cestou treliisom.

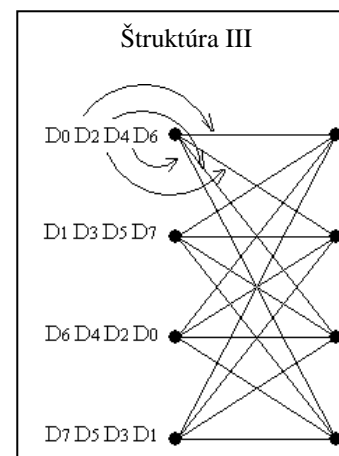
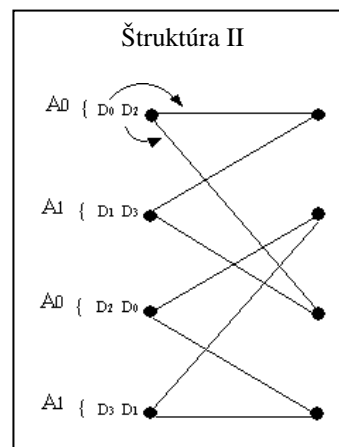
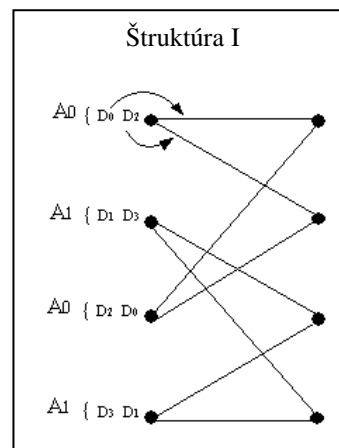
Každý vetve treliisa bola priradená jedna podmnožina kódovej knihy, tak ako je to znázornené na obr. 2. Do každého uzla treliisa musia vstupovať vetvy s najvzdialenejšími podsúbormi t. j. v jednom uzle môžu byť podmnožiny  $D_0, D_2$  a v druhom  $D_1, D_3$ . Každý podsúbor sa v treliise musí vyskytovať minimálne dva krát. V každom podsúbore sa určia euklidovské vzdialenosti vstupného vektora od každého kvantizačného vektora v danom podsúbore a vyberú sa iba vektory s najmenším skreslením. Ďalej, v každom časovom intervale sa z dvoch ciest v každom uzle vyberie cesta s menším skreslením (červená farba na obr.2). Ak počet načítaných vstupných vektorov v pamäti je  $h_p + r$ , vypúšťajú sa  $r$  okvantovaných vektorov z kvantizátora. V tomto kroku sa musia z pamäte vymazať tie cesty treliisa, ktoré v čase  $t - h_p$  smerujú z iného uzla (stavu) ako vybraná cesta s najmenšou akumulovanou metrikou (súčet diečích skreslení, t.j. čísla pri každom uzle treliisa na obr.2). Kódovanie prebieha súčasne s kvantovaním. Z dvoch vetiev vystupujúcich z jedného uzla je vrchnej vetve priradený kanálový symbol 0, spodnej vetve 1 (t.j. určenie podsúboru). Ďalší bit je potrebný na určenie toho, ktorý z dvojice kvantizačných vektorov  $\bar{b}_i$  v konkrétnom podsúbore  $D_i$  bol vybratý.

IV. EXPERIMENTÁLNE VÝSLEDKY SIMULÁCIE A ICH ZHODNOTENIE

Program na implementáciu vektorového kvantovania s treliovým kódovaním bol realizovaný v programovom prostredí C++ a bol použitý na kvantovanie 10000 vstupných vektorov s Gaussovým rozdelením pravdepodobnosti pre každý rozmer  $v = 2, 4, 8, 12$  a  $16$ . Vlastnosti vstupnej postupnosti, z ktorej sa vytvárali vektory požadovaných rozmerov, boli: nulová stredná hodnota, jednotková disperzia, počet skalárov 320000. Simulácie boli vykonané pre prenosové rýchlosti  $n = 1$  až  $8 \text{ bit} / \text{vektor}$ .

Tab. 1  
 Pomery  $S/\check{S}$  pre rôzne štruktúry treliu VKTK v závislosti od meniaceho sa počtu bitov/vektor.

$v = 2$					
$n'$ [bit/vektor]	$S/\check{S}$ [dB]				
	2	3	4	5	6
$n$ [bit/vzorka]	1	1,5	2	2,5	3
Štruktúra I	4,92	7,64	10,42	13,23	15,93
Štruktúra II	4,86	7,58	10,38	13,26	16,22
Štruktúra III	x	7,59	10,39	13,28	16,23
OVK	4,43	7,02	9,51	12,38	15,32
$v = 4$					
$n'$ [bit/vektor]	$S/\check{S}$ [dB]				
	2	3	4	5	6
$n$ [bit/vzorka]	0,5	0,75	1	1,25	1,5
Štruktúra I	2,28	3,7	5,08	6,54	7,94
Štruktúra II	2,27	3,68	5,03	6,53	7,99
Štruktúra III	x	3,55	5,06	6,53	7,99
OVK	1,9	3,37	4,64	6,05	7,49
$v = 8$					
$n'$ [bit/vektor]	$S/\check{S}$ [dB]				
	2	3	4	6	8
$n$ [bit/vzorka]	0,25	0,38	0,5	0,75	1
Štruktúra I	1,08	1,76	2,48	3,99	5,72
Štruktúra II	1,07	1,74	2,48	3,97	5,73
Štruktúra III	x	1,75	2,47	3,98	5,72
OVK	0,87	1,51	2,21	3,7	5,39
$v = 12$					
$n$ [bit/vektor]	$S/\check{S}$ [dB]				
	2	3	4	6	8
$n$ [bit/vzorka]	0,17	0,25	0,33	0,5	0,67
Štruktúra I	0,71	1,15	1,61	2,67	3,92
Štruktúra II	0,71	1,15	1,61	2,66	3,91
Štruktúra III	x	1,15	1,61	2,67	3,91
OVK	0,57	0,97	1,43	2,45	3,67
$v = 16$					
$n'$ [bit/vektor]	$S/\check{S}$ [dB]				
	3	4	5	6	8
$n$ [bit/vzorka]	0,19	0,25	0,31	0,38	0,5
Štruktúra I	0,85	1,21	1,6	2,01	3
Štruktúra II	0,85	1,2	1,59	2,01	2,99
Štruktúra III	0,85	1,2	1,59	2,01	3,01
OVK	0,71	1,06	1,43	1,85	2,81



Obr.3 Typy štruktúr 4-stavového treliu (I a II pre parameter  $M=1$ ; III pre parameter  $M=2$ ).

Parametre vektorového kvantizátora s trellisovým kódovaním: možný počet stavov  $S = 4, 8, 16, 32$  a  $64$ ; rôzne veľkosti hĺbky prehľadávania  $h_p$  (do 100) a periódy vypúšťania  $r$  (tiež do 100). Používateľ najprv zvolí rozmer vektorov kódovej knihy ( $v$ ) a následne jej vektory sú rozdelené do podsúborov  $(D_0, D_1, D_2, D_3)$  pre prípad dvojvetvového trellisa ( $M=1$ ) resp. do ôsmich podsúborov  $(D_0, D_1, D_2, D_3, D_4, D_5, D_6, D_7)$  pre prípad štvorvetvového trellisa ( $M=2$ ).

Z množstva získaných experimentálnych výsledkov sú v tomto článku uvedené iba tie, v ktorých sa testoval vplyv štruktúr trellisa na dosiahnutie maximálnych hodnôt pomeru signálu k šumu ( $S/\bar{S}$ ) pri postupne sa meniacom počte bitov na vektor ( $n$ ). Pomer  $S/\bar{S}$  sa vypočítaval ako pomer výkonu signálu k výkonu chybového signálu v logaritmickú mierku. Ukážka testovaných štruktúr pre prípad 4-stavového trellisa je zobrazená na obr.3. Dosiahnuté výsledky  $S/\bar{S}$  simulácie VKTK pre nekorelované vstupné postupnosti s Gausovým rozdelením pravdepodobnosti sú zapísané v Tab.1, pričom sa menil rozmer euklidovského priestoru ( $v$ ). Parametre VKTK boli: hĺbka prehľadávania  $h_p = 25$ , perióda vypúšťania  $r = 10$ , počet vstupných vektorov 10 000 a počet stavov trellisa  $S = 8$ . Pre porovnanie sú v tabuľke uvedené aj hodnoty pomeru  $S/\bar{S}$  optimálneho vektorového kvantizátora (OVK). Z Tab.1 plynú nasledujúce skutočnosti:

- s rastúcim rozmerom  $v$  rastie pomer  $S/\bar{S}$  pri rovnakej prenosovej rýchlosti  $n$ . Zároveň tento pomer rastie s rastúcou prenosovou rýchlosťou  $n$  pri nemeniacom sa rozmere (čo potvrdilo správnosť navrhnutého algoritmu simulácie)..
- Dosahovaný pomer  $S/\bar{S}$  pri rovnakých hodnotách počtu bitov na vzorku ( $n$ ) a rôznych štruktúrach trellisa je takmer rovnaký, mení sa nanajvýš v desatinách či stotínach dB. Z toho vyplýva, že typ štruktúry trellisa nemá vplyv na výraznú zmenu pomeru  $S/\bar{S}$ , pretože dôležitejšiu úlohu tu zohráva predovšetkým dobre vykonané rozdelenie kódovej knihy kvantizátora a priradenie jej podsúborov jednotlivým uzlom trellisa.
- Dosahovaný pomer  $S/\bar{S}$  simulovaného VKTK je väčší ako pri OVK, čo je spôsobené tým, že VKTK využíva kódovú knihu dvojnásobnej veľkosti (tá sa však v procese trellisového kódovania upraví na požadovanú veľkosť) a tiež tým, že sa využíva pamäť na vybratie optimálnej cesty trellisom VKTK s najmenším celkovým kvantizačným skreslením.
- S rastúcim rozmerom rastie aj počet vektorov kódovej knihy kvantizátora, tým však narastá aj výpočtová náročnosť rozdelenia tejto kódovej knihy.

## V. ZÁVER

Dosiahnuté výsledky pri simulácii vektorového kvantizátora s trellisovým kódovaním ukazujú, že použitie takéhoto kvantizátora je veľmi účinné v porovnaní s optimálnym vektorovým kvantizátorom, ktorého zložitosť návrhu a zložitosť výberu kvantizačných vektorov rastie exponenciálne s prenosovou rýchlosťou údajov a s rozmerom vektora. Vektorové kvantizátory s trellisovým kódovaním ponúkajú redukciu zložitosti v procese vektorového kvantovania. Overenie tohto kvantizátora aj na skúšobných obrazoch potvrdzuje správnosť navrhnutého algoritmu, ako aj správnosť jeho programovej realizácie. Môže byť použitý ako vhodný spôsob kvantovania za účelom zvýšenia výkonu. Tým je možné získať efektívny systém kódovania, ktorý môže byť dôležitou súčasťou digitálneho spracovania signálov z inteligentných senzorov v oblasti elektroniky a komunikačných systémov.

## LITERATÚRA

- [1] Marcellin, M. W. – Fischer, T. R.: *Trellis coded quantization of memoryless and Gauss-Markov sources*. IEEE Trans. Commun., Vol.38, pp.82-93, 1990.
- [2] Fischer, T. R. - Marcellin, M. W. – Wang, M.: *Trellis – coded vector quantization*. IEEE Trans. Inform. Theory, Vol. 37, pp. 1551 – 1556, 1991.
- [3] Ungerboeck, G.: *Trellis-coded modulation with redundant signal sets, Part I and II*. IEEE Commun. Mag., Vol.25, pp.5-21, 1987.
- [4] Gladišová, I.- Mihalík, J.: *Trellis Coding for Vector Quantization*. Proc. 5 th Int. Scient. Conf. "Digital Signal Processing and Multimedia Communications", TU Košice, pp. 31-33, 2001.
- [5] Wang, H. S. – Moayeri, N.: *Trellis coded vector quantization*. IEEE Trans. on Commun., Vol.40, No.8, pp.1273-1276, 1992.
- [6] Freeman, G. H. – Blake, I. F. – Mark, J. W.: *Trellis source code design as an optimization problem*. IEEE Trans. Inform. Theory, Special Issue on Coding : Theory and Techniques, pp. 1– 48, 1988.
- [7] Bilgin, A. – Sementilli, P. J. – Marcellin, M. W.: *Progressive Image Coding Using Trellis Coded Quantization*. IEEE Trans Image Processing, Vol. 8, No. 11, pp. 1638-1643, Nov. 1999.
- [8] Pradhan, S.S. - Ramchandran, K.: *Distributed Source Coding Using Syndromes (DISCUS): Design and Construction*. IEEE Tans. on Inform. Theory, Vol. 49, No. 3, 2003.
- [9] Michalčín, V.: *Subband Vector Quantization with Trellis Coding*. Acta Electrotech. et Informatica, Vol.11, No.1, pp. 48-53, 2011.



# Výber kandidátskych uzlov na báze teórie hier v hybridných MANET - DTN

<sup>1</sup>Ján Papaj, Tomáš Loziak

<sup>1</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>jan.papaj@tuke.sk

**Abstrakt** — Smerovacie protokoly Mobilných Ad-hoc sietí (MANET) boli navrhnuté za účelom vytvorenia viacsokovej cesty medzi dvojicou komunikujúcich uzlov. Tieto mobilné uzly sú schopné prevádzky aplikácií na MANET prístupe. V kontraste s týmto prístupom, komunikačné schémy založené na prístupe Delay Tolerant Networks (DTN) implementujú asynchrónnu komunikáciu (často vyžadujú použitie nových aplikácií). Výhodou sietí DTN je vyššia doručiteľnosť dát i za cenu vyšších hodnôt oneskorenia prevažne v riedko osídlených oblastiach. V tomto článku je navrhnutá hybridná schéma, ktorá kombinuje DSR protokol s výberom kandidátskeho uzla, ktorý je založený na prístupe DTN s cieľom využitia výhod oboch prístupov. Výber kandidátskych uzlov je uskutočnený na základe smerovacích parametrov DSR použitím Hry Parametrov a kladie dôraz na bezpečnosť vo výbere ďalšieho preskoku.

**Keywords** — DTN, Kandidátsky uzol, MANET, teória hier.

## I. ÚVOD

S rozvojom mobilných komunikačných sietí rastie najmä potreba zabezpečenia stále väčšej mobility používateľa. S tým súvisí aj potreba fungovania mobilných sietí v núdzových a krízových prípadoch. Myslia sa tým najmä prípady, keď pevné a mobilné siete s infraštruktúrou nie sú k dispozícii (napr. sú zničené alebo preťažené) a komunikačné siete musia pracovať v skutočne náročných prostrediach. Takéto situácie môžu vykazovať rôzne vlastnosti konektivity a s nimi späté problémy, najmä pokiaľ ide o smerovanie. Smerovacie protokoly určené pre tieto mimoriadne okolnosti musia rátať s chybami prenosu, zlyhaním mobilného uzla a dokonca aj so segmentovaním siete na viaceré časti bez vzájomného mobilného dosahu. Z týchto požiadaviek vyplynula potreba vytvorenia takých smerovacích algoritmov, ktoré v sebe zahrňujú mechanizmy známe z Oportunistických sietí (DTN) - aby bolo možné vyrovnáť sa s roztrhnutím mobilnej siete, a zároveň mechanizmy zo smerovacích protokolov MANET aby bola zachovaná vyššia efektívnosť prenosu prostredníctvom sieťových zdrojov (Fig. 1) [1].



Obr. 1 Rozdiel medzi MANET a hybridnou MANET-DTN



### A. Smerovanie v hybridných MANET – DTN sieťach

Smerovacie protokoly v hybridných MANET-DTN sieťach môžeme rozdeliť podľa troch základných prístupov (Tabuľka 1) [1], [7] :

a) *Integrácia DTN mechanizmov do MANET protokolov* - tento typ prístupu integruje DTN mechanizmy do tradičných MANET protokolov. Najčastejšie tento prístup mení už existujúci protokol MANET, ktorý zabezpečuje prevádzku existujúceho spojenia koncových zariadení, aby zabezpečoval komunikáciu aj medzi rozdelenými segmentmi siete. K dosiahnutiu takejto komunikácie je potrebné integrovať prinajmenšom mechanizmus *store-and-forward*. Výberom optimálnych uzlov na základe predikcie budúcich komunikačných príležitostí a replikovaním správ v sieti, môžeme zvýšiť celkovú smerovacuú výkonnosť algoritmov. Prístupy SF-BATMAN[2] a STAR[3] využívajú iba použitie vyrovnávacej pamäte. DT-DYMO[4] a CAR [5] dodatočne využívajú modifikované kontrolné smerovacie správy na pravdepodobnostný výber smerovacích uzlov. Aby došlo k ušetreniu sieťových zdrojov, žiadny z týchto protokolov nereplikuje pakety. Replikovanie správ používa napríklad prístup HYMAD[6]. Výhoda takéhoto druhu prístupov spočíva v možnosti ich implementácie na sieťovej vrstve bez potreby zásahu do vyšších vrstiev. Avšak z pohľadu aplikácií nie sú vhodné pre určité typy. Multimediálne aplikácie v reálnom čase nefungujú ako v sieťach s infraštruktúrou, pretože mechanizmus ukladania paketov vykazuje vyššie oneskorenie a kolísanie oneskorenia.

b) *Podpora DTN bundle protokolu* - prístupy tohto druhu kombinujú DTN bundle protokol s tradičným MANET smerovaním. *Bundle protokol* je protokol aplikačnej vrstvy fungujúci na princípe *store-and-forward* [8]. Základná jednotka tohto protokolu je časť bundle, ktorá má väčšiu kapacitu ako dátové jednotky nižších vrstiev, čiže segmenty a pakety. Napriek tomu, že bundle protokol nedefinuje ako majú byť jednotlivé bundle smerované, DTN smerovací protokol ich smeruje pri prenose dát. Okrem iného rozhoduje aj o tom, či budú pakety preposielané a ako sa vyberú kandidátske uzly (najčastejšie sa vyberajú na základe pravdepodobnosti). Z uvedeného dôvodu je výkonnosť týchto prístupov najviac ovplyvňovaná použitým DTN smerovacím mechanizmom [9]. Pretože bundle protokol je integrovaný na aplikačnej vrstve, rozhodnutie či využiť MANET alebo DTN smerovanie musí prebehnúť tiež na aplikačnej vrstve.

c) *Návrh nového smerovacieho protokolu pre rozličné siete* - vyskytujú sa tiež prístupy, ktoré nekombinujú alebo neintegrujú existujúce MANET a DTN protokoly, ale sú špeciálne vyvíjané pre široké spektrum sietí s rozličnými vlastnosťami konektivity. Kľúčom k úspešnému vývoju takéhoto hybridného protokolu je nájdenie smerovacej metriky, ktorá funguje pod rozličnými podmienkami konektivity. Hlavnou výhodou takýchto protokolov je, že sú navrhnuté presne pre hybridné siete a využívajú mechanizmy a metriky vhodné pre spomínané siete. Na druhej strane návrh a implementácia úplne nového protokolu je náročný proces [10].

Tabuľka 1  
 Prehľad súčasných smerovacích protokolov v hybridných MANET-DTN sieťach

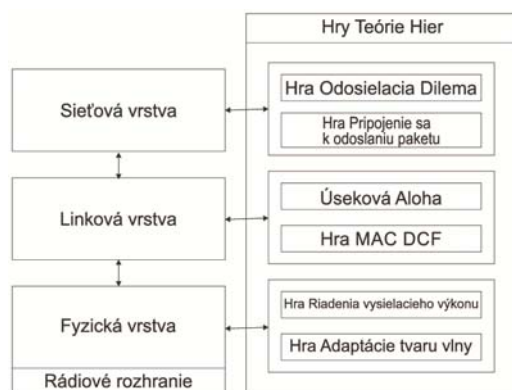
Protokol	MANET protokol (metrika)	Bundle	Preposielanie správy	Výber kandidátskeho uzla
AODV-DTN	AODV	Voliteľný	Voliteľný	Voliteľný
BATMAN-SF	BATMAN	Nie	Nie	Nie
CAR	DSDV	Nie	Nie	Áno
DT-DYMO	DYMO	Nie	Nie	Áno
DTS-OLSR	OLSR	Áno	Voliteľný	Voliteľný
HYMAD	Vektor vzdialeností	Nie	Áno	Áno
Lakkakorpi a spol.	AODV	Voliteľný	Voliteľný	Voliteľný
R3	Očakávané oneskorenie	Nie	Áno	Nie
STAR	OLSR	Nie	Nie	Nie

### B. Teória hier v hybridných MANET – DTN sieťach

Teória hier (TH) je vedná disciplína [11], ktorá sa zaoberá štúdiom a analýzou racionálneho správania sa jednotlivcov a ich vzájomnej interakcie. Skúma najmä spôsoby ako strategické interakcie medzi subjektmi vytvárajú výsledky – napríklad úžitok. Jedná sa prevažne o štúdium matematických metód interakcie a spolupráce medzi subjektmi, ktoré sa rozhodujú racionálne. Ich rozhodnutie pritom ovplyvní stav toho druhého. Objektom skúmania TH je pochopenie, vysvetlenie a snaha predpovede možného výsledku tejto interakcie. V podstate ide o predvídanie výsledkov vlastného konania na základe rozhodnutia ostatných činiteľov. Pri takomto postupe predpokladáme, že aj títo činitelia sa rozhodujú racionálne (Obr. 2).

Ak rozprávame o vedomej a cieľovej interakcii jednotlivca s okolitým prostredím, v TH sa

používa pojem hra. Hra musí mať vopred stanovené pravidlá, pri rôznych výstupoch má rôzne hodnoty, ktoré ovplyvňujú hráči svojim správaním [11]. O hre rozprávame iba vtedy, ak existujú viaceré možnosti výsledných stavov. Ak rôzne rozhodnutia hráča povedú k stále rovnakému výsledku, potom sú tieto rozhodnutia nepodstatné a v takomto prípade sa nejedná o hru. Ďalší veľmi dôležitým pojmom sú hráči v hre. Hráči sú spomínané subjekty, ktoré sa správajú racionálne. Prevažne sa za hru považuje interakcia hráča k inému hráčovi, alebo k skupine hráčov, ale za hru sa považuje aj situácia interakcie jedného hráča voči systému. Takáto hra sa nazýva hra jedného hráča alebo hra proti prírode. Dôležitá súčasť hry sú ťahy hráča. Ak sa hráč ocitá v neopakovateľnej situácii, takýto ťah sa nazýva voľba. V opačnom prípade, ak ide o dlhodobjšiu činnosť, alebo hráč reaguje na voľbu iných hráčov v predchádzajúcich hrách, hovoríme o stratégii. Výstupy TH sú závislé od použitých volieb a stratégií. Cieľom TH je získanie maximálneho možného úžitku. Ak sú výstupy hráča v zhode s jeho zámermi, hovoríme o jeho výhre, zisku alebo výplate. Pri opačnej situácii – ak sú výstupy hráča v rozpore s jeho zámermi, hovoríme a prehre resp. strate.



Obr. 2 Najznámejšie hry aplikované v bezdrôtových sieťach.

Teória hier je relatívne mladá vedná disciplína [12]. Jej začiatky možno datovať od prvej polovice 20. storočia. Za jej zakladateľa sa považuje John von Neumann, ktorého práca našla uplatnenie najmä v ekonomickej oblasti. Odvtedy sa však jej použitie rozšírilo do mnohých ďalších oblastí. Medzi najpoužívanejšie oblasti patrí ekonómia, informatika, politické vedy, právo, biológia, filozofia. V informatike sa využíva v distribuovanom programovaní, algoritmickej analýze ako aj v počítačových sieťach či už pevných, alebo bezdrôtových. V počítačových sieťach je možné jej použitie pri riadení výkonu, adaptívnom vyhýbaní interferencii, formulácii siete, smerovaní, riadení toku, alebo pri riadení preťaženia liniek.

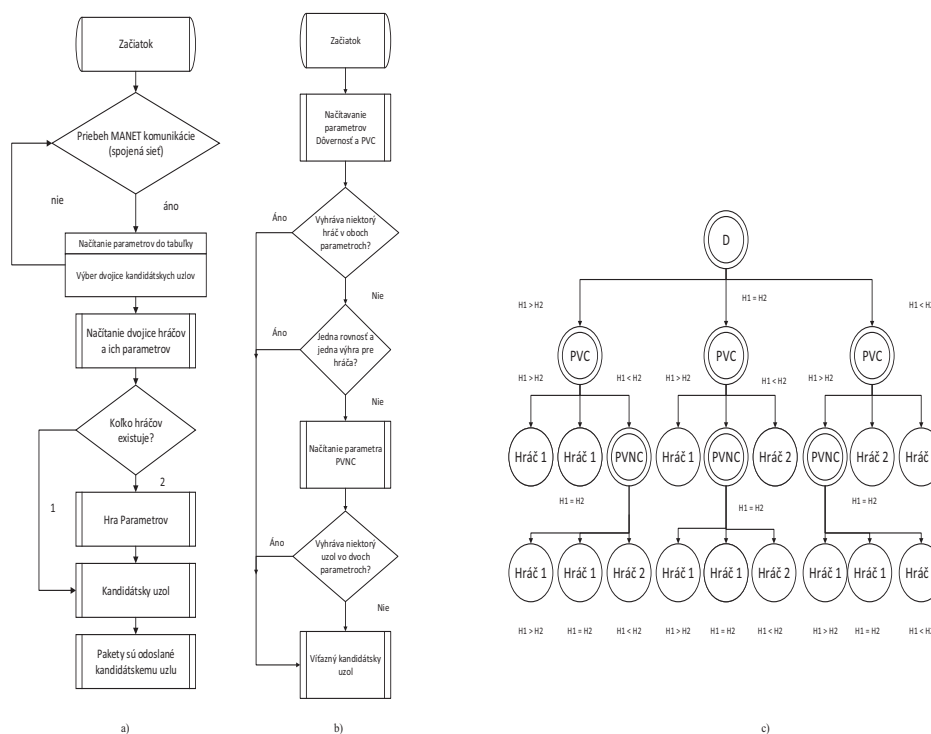
Množinu hráčov predstavujú prevažne mobilné uzly v sieti. Pri špecifických situáciách môže hráča predstavovať aj statické zariadenie väčšinou na okraji siete slúžiace na spojenie s okolitým svetom, nazývané aj brána siete. Množinu ťahov môže obsahovať rôzne možnosti v závislosti od konkrétnej hry. Najpoužívanejšie ťahy sú vysielanie, načúvanie, odoslanie paketu, príjem paketu, ale aj zvolenie modulačnej schémy, kódovej rýchlosti, či úrovne vysielacieho výkonu. V hrách, ktoré sa zaoberajú bezpečnosťou sa vyskytujú aj ťahy ako útočiť, brániť sa, načúvať. Množina ziskov rovnako závisí od typu hry. Najčastejšie ušetrenie energie, frekvenčného pásma, zníženie oneskorenia, SNR, alebo zvýšenie priepustnosti siete, či prenosovej rýchlosti.

## II. NÁVRH ALGORITMU VÝBERU KANDIDÁTSKYCH UZLOV NA BÁZE TEÓRIE HIER V HYBRIDNEJ MANET-DTN SIETI

Navrhnutý algoritmus je založený na monitorovaní parametrov, ktoré sa následne aplikujú do procesu výberu kandidátskych uzlov. Algoritmus výberu kandidátskeho uzla je zobrazený na Obr.3a. Na začiatku keď je sieť nerozpojená a prebieha komunikácia nemodifikovaným smerovacím protokolom DSR, návrh sleduje tri parametre po celú dobu až do rozpojenia end-to-end spojenia:

- *Posledné dva uzly v sieti, ktoré odoslali kontrolnú správu RERR o nedosiahnuteľnosti cieľového uzla.*
- *Koľkokrát sa uzly v sieti vyskytovali v smerovacej ceste k cieľovému uzlu (PVC – Počet Výskytov v Ceste).*

- *Koľkokrát sa uzly vyskytovali v náhradných cestách v smerovacej pamäti (PVNC –Počet Výskytov v Náhradnej Ceste).*
- *Parameter Dôvernost' (D - Dôvernost').*



Obr. 3 Navrhnutý algoritmus pre hybridnú MANET-DTN: a) Algoritmus výberu kandidátskeho uzla, b) Hra Porovnanie parametrov, c) Definovanie výstupných stavov

Ak toto spojenie vypadne, vyberú sa dvaja hráči do teórie hier medzi ktorými bude prebiehať Hra Parametrov. Hra porovnanie parametrov slúži na to aby z dvoch kandidátskych uzlov vybrala jeden, ktorý je vhodnejší na základe získaných parametrov. Algoritmus hry je zobrazený na Obr.3b. Ako je vidieť na vývojovom diagrame, Hra porovnanie parametrov načíta najprv dva primárne parametre D a VPC. Ak má jeden z uzlov lepšie hodnoty v oboch týchto primárnych parametroch stáva sa víťazom. Jeden z uzlov víťazi aj vtedy, ak sa mu jeden primárny parameter zhoduje, ale v druhom má výhodnejšie hodnoty. V prípade, že oba uzly víťazia v jednom primárnom parametre, načíta sa sekundárny parameter PVNC, ktorý rozhodne o víťaznom kandidátskom uzle. Ak sa hráči zhodujú aj v sekundárnom parametre, víťazom je prvý hráč, pretože od neho prišla správa RERR ako posledná. Nasledujúci stromový diagram zobrazuje všetky možné stavy v hre a jeho víťazný kandidátsky uzol (Obr. 3c).

### III. SIMULÁCIE A VÝSLEDKY

OPNET modeler je profesionálny simulačný nástroj, ktorý je využitelný na testovanie navrhnutého mechanizmu výberu kandidátskych uzlov v hybridnej MANET-DTN sieti. OPNET ponúka mnoho užitočných nástrojov pre analýzu mobilných sietí a umožňuje simulovať MANET siete, ktoré využívajú rozličné smerovacie protokoly [13]. Z pohľadu analyzovania funkcionality navrhnutého riešenia je protokol DSR (Dynamic Source Routing) chápaný ako referenčná hodnota pre analyzovanie správania MANET [14]. Za účelom overenia funkčnosti hybridnej MANET - DTN, sme implementovali mechanizmus PROPHET do simulátora OPNET modeler. Počas simulácií sme analyzovali tieto modely:

- *DSR* – model MANET siete s reaktívnym protokolom DSR, ktorý nemá aplikovaný mechanizmus na prenos dát v rozpojenej sieti DTN ani mechanizmus výberu kandidátskych uzlov [14].
- *MANET\_TH* – model MANET siete s implantovaným našim návrhom výberu kandidátskych uzlov.

Simulačné parametre použité v simulačnom nástroji OPNET modeler sú znázornené

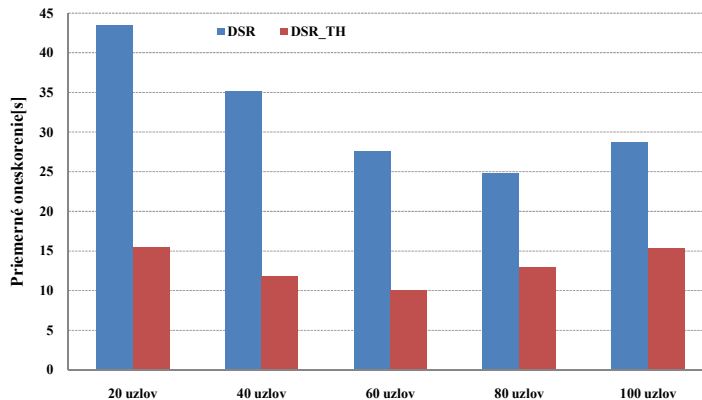
v Tabuľke 2. Počas simulácií sme analyzovali tieto parametre:

- *Priemerná hodnota oneskorenia* - definuje priemernú hodnotu času potrebného na prenos dát medzi zdrojovým a cieľovým uzlom
- *Priemerný počet skokov* - definuje priemerný počet preskokov potrebných na dosiahnutie cieľového uzla
- *Priemerný počet odoslaných paketov* - definuje priemerný počet paketov odoslaných medzi zdrojovým a cieľovým uzlom
- *Priemerný počet prijatých paketov* - definuje priemerný počet paketov prijatých medzi zdrojovým a cieľovým uzlom

Tabuľka 2  
 OPNET modeler simulačné parametre

Parametre	Hodnoty
Smerovacie protokoly	DSR, DSR_TH, PRoPHET
Simulovaná oblasť	2000 m x 2000 m
Čas simulácie	600 s
Dosah mobilných uzlov	250 m
Počet uzlov	20, 40, 60, 80, 100
Rýchlosť uzlov	0 - 6 m/s
Rozmiestnenie uzlov	Náhodné
Model mobility	Náhodný model pohybu
Vysielač výkon	1 mW
Počet hodnôt počas simulácie	1000
Počet opakovaní	10

Počas experimentov sme sa zamerali na situáciu, keď sú komunikačné cesty medzi susednými uzlami rozpojené na krátky čas. To znamená, že je aktivovaný algoritmus výberu kandidátskych uzlov na prenos dát. Rozpojenie komunikačných ciest je simulované modelom náhodného pohybu ako aj krátkym dosahom jednotlivých mobilných uzlov. Rozpojenia predstavujú čas, počas ktorého smerovací protokol DSR nie je schopný nájsť komunikačnú cestu medzi zdrojovým a cieľovým uzlom, resp. medzi susednými uzlami.



Obr. 4 Priemerné oneskorenie doručovania paketov pre hybridnú MANET-DTN sieť.

Na Obr. 4 sú znázornené výsledky pre parameter Priemerné oneskorenie doručovania paketov. Graf predstavuje závislosť tohto parametra na počte uzlov pohybujúcich sa náhodným pohybom. Ak sa komunikačné cesty medzi mobilnými terminálmi rozpoja, smerovací protokol DSR nie je schopný okamžite nájsť adekvátnu cestu, čo sa prejaví zvýšenou hodnotou oneskorenia.

Na druhej strane, implementovaný algoritmus umožní aktivovanie protokolu PRoPHET a vyberie optimálny kandidátsky uzlov na prenos dát. Zo získaných výsledkov vyplýva, že algoritmu výberu kandidátskych uzlov, v porovnaní so štandardným protokolom DSR, predstavuje zníženie tohto parametra takmer o 50%.

V Tabuľke 3 sú uvedené priemerné hodnoty odoslaných aj prijatých smerovacích dát. Z tabuľky vyplýva, že implementácia algoritmu výberu kandidátskych uzlov v prípade rozpojenia komunikačných ciest, predstavuje zvýšenie týchto priemerných hodnôt priemerných je vo všeobecnosti vyššia ako v prípade DSR. Je to spôsobené tým, že jeden z rozhodujúcich faktorov vstupujúcich do výberu kandidátskeho uzla je dôvernosť a preto sa prenáša väčšie množstvo smerovacích dát.

Tabuľka 3  
 Analyzovanie prenosových vlastností pre hybridnú MANET-DTN sieť [kbit/s]: a) priemerný počet odoslaných paketov b) priemerný počet prijatých paketov

a)					b)				
Počet uzlov	Model	Rýchlosť uzlov			Počet uzlov	Model	Rýchlosť uzlov		
		2 m/s	4 m/s	6 m/s			2 m/s	4 m/s	6 m/s
20	DSR	708	752	650	20	DSR	606	1298	768
	DSR TH	700	826	947		DSR TH	740	1514	1838
40	DSR	1039	2003	1185	40	DSR	960	3021	2315
	DSR TH	1989	2399	3204		DSR TH	2014	7546	7614
60	DSR	2061	3298	2439	60	DSR	2044	7256	6701
	DSR TH	2924	2160	4957		DSR TH	3198	20346	16971
80	DSR	4090	5931	3910	80	DSR	4193	9310	7839
	DSR TH	5982	6561	7625		DSR TH	6262	47401	42125
100	DSR	8306	8784	1000	100	DSR	16014	15863	8624
	DSR TH	9122	1139	1269		DSR TH	25367	24787	48396

#### IV. ZÁVER

V tomto príspevku sme predstavili mechanizmus výberu kandidátskych uzlov na báze teórie hier v hybridných MANET-DTN sieťach. Hlavnou ideou bol návrh a implementácia výberu bezpečného kandidátskeho uzla na základe teórie hier. Tento algoritmus umožňuje taktiež vybrať dôveryhodný mobilný uzol, ktorý bude následne prenášať dáta v prípade rozpojených spojení komunikačných ciest. Hybridné MANET – DTN siete poskytujú nové možnosti využitia mobilných sietí bez ohľadu na využiteľnú infraštruktúru.

Predložený článok poskytuje aj analýzu funkčnosti činnosti navrhnutého algoritmu v hybridnej MANET-DTN. Na základe týchto výsledkov môžeme konštatovať, že algoritmus poskytuje účinný mechanizmus výberu uzlov, ktoré následne budú využité na prenos dát medzi mobilnými terminálmi.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### REFERENCIE

- [1] Ch. Raffelsberger, H. Hellwagner, "Overview of Hybrid MANET-DTN Networking and its Potential for Emergency Response. Operations Electronic Communications of the EASST," vol. 56, pp.1-12, 2013.
- [2] L. Delosieres, S. Nadjim-Tehrani, "BATMAN store-and-forward: The best of the two worlds," In Proc. PerComWorkshops '12. Pp. 721–727. IEEE, 2012.
- [3] S. Jain, S. Gopinath, D. Raychaudhuri. STAR, "Storage aware routing protocol for generalized Delay Tolerant Networks," In Proc. WoWMoM '11. pp. 1-4. IEEE, 2011.
- [4] C. Kretschmer, S. Ruhup, C. Schindelbauer, "DT-DYMO: Delay-Tolerant Dynamic MANET On-demand Routing," In Proc. ICDCSW '09. Pp. 493–498. IEEE, 2009.
- [5] M. Musolesi, S. Hailes, C. Mascolo, "Adaptive routing for intermittently connected mobile ad hoc networks," In Proc. WoWMoM '05. Pp. 183–189. IEEE, 2005.
- [6] J. Whitbeck, V. Conan, *HYMAD: Hybrid DTN-MANET routing for dense and highly dynamic wireless networks*, Computer Communications, Vol. 33, No.13, pp.1483–1492, 2010.
- [7] J. Ott, D. Kutscher, C. Dwertmann, "Integrating DTN and MANET routing," In Proc.CHANTS '06. pp. 221–228. ACM, 2006.
- [8] J. Lakkakorpi, M. Pitkanen, J. Ott, "Adaptive routing in mobile opportunistic networks," In Proc. MSWIM '10. pp. 101–109. ACM, 2010.
- [9] R. Pant, A. Tunpan, P. Mekbungwan, R. Virochpoka, K. Kanchanasut. DTN overlayon OLSR network. In Proc. AINTEC '10. Pp. 56–63. ACM, 2010.
- [10] X. Tie, A. Venkataramani, A. Balasubramanian, "R3: Robust replication routing in wireless networks with diverse connectivity characteristics," In Proc. MobiCom '11. pp. 181–192. ACM, 2011.
- [11] A. Démuth, "Teória hier a problém rozhodovania", Vysokoškolský učebný text, 2013. ISBN: 978-80-8082-580-5
- [12] S. Mehta and K. S. Kwak (2010), *Application of Game Theory to Wireless Networks, Convergence and Hybrid Information Technologies*, Marius Crisan (Ed.), ISBN: 978-953-307-068-1, InTech, DOI: 10.5772/9642.
- [13] OPNET Modeler Simulation Software, <http://www.opnet.com>.
- [14] A. Čížmár, E. Doboš, J. Papaj, "Security and QoS Integration Model for MANETs," In Computing and Informatics Vol. 31, No.5, pp. 1025-1044, ISSN: 1335-9150, 2012.

# Využitie dôvernosti pri výbere kandidátskych uzlov v hybridnej MANET-DTN

<sup>1</sup>Ján Papaj, Peter Malic

<sup>1</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>jan.papaj@tuke.sk

**Abstrakt** — Tento článok je zameraný na využitie dôvernosti pri výbere kandidátskych uzlov v hybridných MANET-DTN sieťach. Kandidátsky uzol je vybraný uzol, ktorý sa bude pokúšať o prenesenie dát medzi izolovanými skupinami mobilných uzlov. Dôraz článku je kladený na dôvernosť ako bezpečnostný koncept používaný pri návrhoch hybridných MANET - DTN sietí. Navrhnutý algoritmus umožňuje vyberať také uzly, ktoré sa z pohľadu činnosti javia ako dôverné a teda môžu byť využívané ako prenosové médium. Tento algoritmus je implementovaný do prostredia simulátora OPNET modeler, kde je vykonaná aj analýza funkčnosti činnosti.

**Keywords** — Bezpečnosť, dôvernosť, kandidátsky uzol, hybridné MANET-DTN siete.

## I. ÚVOD

S výhodami spojenými s mobilnými zariadeniami, mnohými scenármi sietí a aplikáciami ľudia začínajú čoraz viac využívať mobilitu a vzájomnú kooperáciu. V otvorenom, voľne prístupnom bezdrôtovom prostredí je jednou z hlavných požiadaviek bezpečnosť v sieti. Základným prvkom každej bezpečnosti, ktorá je zameraná na množstvo používateľov je dôvernosť. V decentralizovanom a rozloženom prostredí, kde nejestvuje žiadna štruktúra je dôvernosť ešte vyššou prioritou. Vypočítavaná dôvernosť je bezpečnostný koncept a je alternatívou kryptografie, rôznych bezpečnostných certifikátov a digitálnych podpisov s tým rozdielom, že nepotrebuje dôveryhodnú autoritu. Umožňuje uzlom v sieti predpokladať budúce správanie iných uzlov. Najnovším trendom je integrácia dôvernosti do sietí typu hybridné MANET-DTN. Sú to hybridné siete využívajúce výhody mobilných Ad-Hoc sietí a sietí tolerujúcich oneskorenie v doručení informácií, DTN (Delay Tolerant Networks).

### A. Mobilné Ad-Hoc siete (MANET)

MANET sú veľmi flexibilné siete. Minimálna konfigurácia staníc umožní ich rýchle nasadenie. Môžu byť teoreticky postavené kdekoľvek použitím bezdrôtového zariadenia, ktoré podporuje spojenie typu end-to-end. Sú vhodným komunikačným prostriedkom v núdzových situáciách, ako sú prírodné katastrofy, ozbrojené konflikty, na miestach s chýbajúcou infraštruktúrou. Aby mohli uzly v takejto sieti smerovať dáta naprieč celou sieťou, musia byť použité smerovacie protokoly. Do týchto smerovacích protokolov sa momentálne integruje bezpečnostný koncept dôvernosti medzi uzlami. Sieť typu MANET je autonómny systém, ktorý nevyžaduje pripojenie do infraštruktúry. Hlavné charakteristiky sietí MANET môžeme zhrnúť do týchto bodov [1]:

- *Dynamická topológia* - topológia sa dynamicky mení, pretože uzly sú mobilné terminály a komunikácia prebieha na vzdialenosť viacerých preskokov.
- *Obmedzená šírka pásma, variabilná kapacita linky* - šírka pásma a prenosová rýchlosť v týchto sieťach závisí od použitého komunikačného štandardu. Predpokladá sa, že uzly v sieti nie sú rovnaké a každý z nich má rôzne hardvérové vlastnosti.
- *Energeticky obmedzené operácie* - mobilné uzly sú závislé na obmedzených zdrojoch energie, preto aj operácie a funkcie by mali brať ohľad na túto základnú vlastnosť mobilných terminálov.

- *Limitovaná fyzická bezpečnosť* - mobilné siete sú väčšinou viac náchylné na fyzické útoky ako siete s pevnou infraštruktúrou. Preto by mala byť podrobne zvážená hrozba útokov ako sú napríklad DoS (Denial of Service attack), odpočúvanie atď.

## B. Siete typu DTN

Tieto siete boli pôvodne navrhnuté pre siete, ktoré spájajú obrovské vzdialenosti. Sú navrhnuté tak, aby bolo tolerované a možné aj veľké oneskorenie pri doručovaní správ. Tento typ sietí sa ujal a používa sa aj pri senzorových systémoch s občasnou konektivitou, satelitných sieťach s periodickou konektivitou, ďalej pri podmorských senzorových akustických sieťach s oneskorením a častým prerušovaním spojenia, ktoré je spôsobené vplyvmi prostredia. V tomto druhu sietí je definovaná medzivrstva nazývaná bundle layer. Nachádza sa nad sieťovou vrstvou z pohľadu referenčného ISO OSI modelu. Siete typu DTN boli navrhnuté preto, lebo bežný Internet protokol je striktné navrhnutý a nepracuje správne v určitých podmienkach. Jedným zo spôsobov použitých pre komunikáciu v DTN sieťach je integrovanie tzv. bundle vrstvy do sieťového modelu. Táto vrstva plní nasledovné funkcie:

- Vrstva bundle*: Aplikácie, ktoré podporujú DTN prenos dát posielajú správy ľubovoľnej dĺžky nazývané ADU (Application Data Units) [2], [3] Tieto dátové jednotky sú väčšinou prenášané a kompletne doručené na aplikačnú vrstvu. ADU sú transformované počas prenosu vrstvou bundle na protokolové dátové jednotky nazývané zväzky (bundles). Tieto zväzky sú potom prenášané pomocou DTN uzlov.
- DTN uzly*: DTN uzol môžeme vo všeobecnosti označiť ako zariadenie, ktoré umožňuje preposielanie a prijímanie zväzkov. Uzly majú implementovanú podporu pre vrstvu bundle [2]. Uzly môžu byť členmi skupiny koncových zariadení označovanej ako DTN Endpoints. Toto spájanie uzlov do skupín sa používa pre posielanie unicastovej, multicastovej alebo broadcastovej správy.
- Bezpečnosť vo vrstve bundle*: V tejto novej vrstve je dôležitá otázka bezpečnosti. V prostredia, pre ktoré sú siete DTN navrhnuté, otázka neautorizovaného prístupu a zneužitia tejto vrstvy veľmi dôležitá. Preto je potrebné navrhnuť nové metódy a mechanizmy na poskytovanie bezpečnosti.

## II. DÔVERNOSŤ V HYBRIDNÝCH MANET-DTN

Dôvernosť pre určitý uzol v sieti je subjektívne posúdenie agenta, respektíve ostatných uzlov na spoľahlivosť a správnosť prijímaných, odoslaných informácií cez tento uzol v danom kontexte. Dôvernosť odráža vieru alebo očakávanie na spoľahlivosť, dostupnosť, integritu a kvalitu služby cieľového uzla z pohľadu budúcej aktivity a správania sa tohto uzla [4].

### A. Metrika a parametre dôvernosti

Na základe rôznych vstupných parametrov určujeme metriku. Na základe tejto metriky sme schopní určiť hodnotu dôvernosti. Tieto metriky môžeme zaradiť do troch základných kategórií [4]:

- Trust scale* - v tomto ponímaní má vypočítaná dôvernosť diskkrétne, alebo spojité hodnoty v určitom rozsahu a je použitý vopred určený prah pre porovnanie a vyhodnotenie.
- Trust facets* - v tomto modeli sú definované *istota* a *dôvernosť* v rovnakom rozsahu napríklad [0; 1] a spolu predstavujú dôveryhodnosť. Túto dôveryhodnosť si môžeme predstaviť ako najkratšiu vzdialenosť od počiatku ku bodu so súradnicami [dôvernosť, spoľahlivosť] v 2D priestore. Metrika môže byť zložená aj z viacerých hodnôt, napr.  $b+d+u=1$ , kde je *b*- dôvera (belief), *u*- neistota (uncertainty) a *d*-nedôvera (disbelief). Dôvernosť je potom reprezentovaná v tomto priestore.
- Trust logics* - niektoré prístupy používajú pravdepodobnosť ako metriku pre určenie dôvernosti. Ako parametre vstupujúce do výpočtu tejto pravdepodobnosti, napríklad použitím Bayesovej štatistiky, sa používajú: počet prijatých odoslaných paketov, počet chybných paketov, počet interakcií, počet úspešných neúspešných pripojení.

### B. Vlastnosti dôvernosti

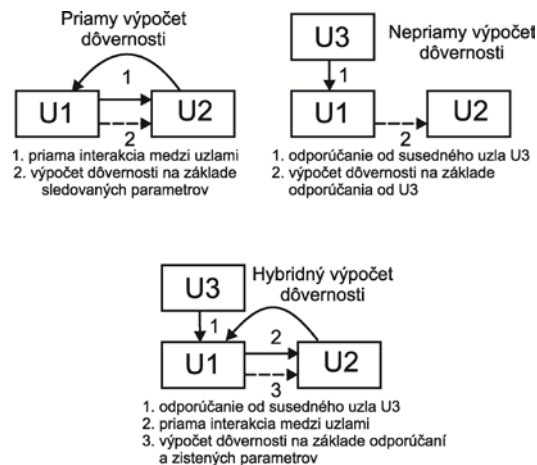
Spomenieme základné vlastnosti dôvernosti, ktoré sú veľmi dôležité pri výpočte. Môžeme hovoriť o troch základných: *asymetria* (asymetry), *tranzitívnosť* (transitivity) a *možnosť skladania* (composability). Asymetria znamená, že ak uzol A dôveruje uzlu B s určitou hodnotou dôvernosti, nie je isté, že aj uzol B dôveruje uzlu A tou istou hodnotou. Tranzitívnosť znamená, že dôvernosť môže byť širiteľná danou cestou. Ak uzol A dôveruje uzlu B a uzol B dôveruje uzlu C znamená to, že uzol A dôveruje uzlu C presnou hodnotou. Možnosť skladania



znamená, že informácia o dôvernosti prijatá prostredníctvom dostupných uzlov môže byť zložená a môže tvoriť jednu hodnotu [5].

### C. Metódy výpočtu dôvernosti

Výpočet dôvernosti je zložený z troch komponentov: *skúsenosť (experience)*, *odporúčanie (recommendation)* a *vedomosť (knowledge)*. Prvý komponent je priamo zisťovaný pomocou pripojených uzlov, respektíve uzlov v dosahu a je aktualizovaný v regulárnych časových intervaloch a uložený v tabuľke. Táto aktualizovaná tabuľka je odosielaná susedným uzlom ako *odporúčanie (recommendation)*. V regulárnych intervaloch sú predošlé zistené hodnoty dôvernosti zohľadnené v aktuálnej hodnote *vedomosti (knowledge)* a všetky tieto tri hodnoty sa používajú buď individuálne, alebo ich kombináciou pri výpočte výslednej hodnoty [6], [7].



Obr. 1 Modely výpočtu dôvernosti

Výpočty môžeme klasifikovať do dvoch základných kategórií. Tieto kategórie sa potom ďalej delia:

- **Distribuovaný výpočet dôvernosti (distributed trust computation)** - každý uzol počíta svoje vlastné hodnoty dôvernosti ku uzlom v jeho dosahu. Patria tu (Fig. 1):
  - a) *Priamy výpočet dôvernosti (Direct trust)*
  - b) *Výpočet na základe odporúčaní (Recommendation based- indirect trust)*
  - c) *Hybridná metóda výpočtu*
- **Centralizovaný výpočet dôvernosti (centralized trust computation)** - centralizovaný prvok pomáha a manažuje výpočet v jednotlivých uzloch v sieti.

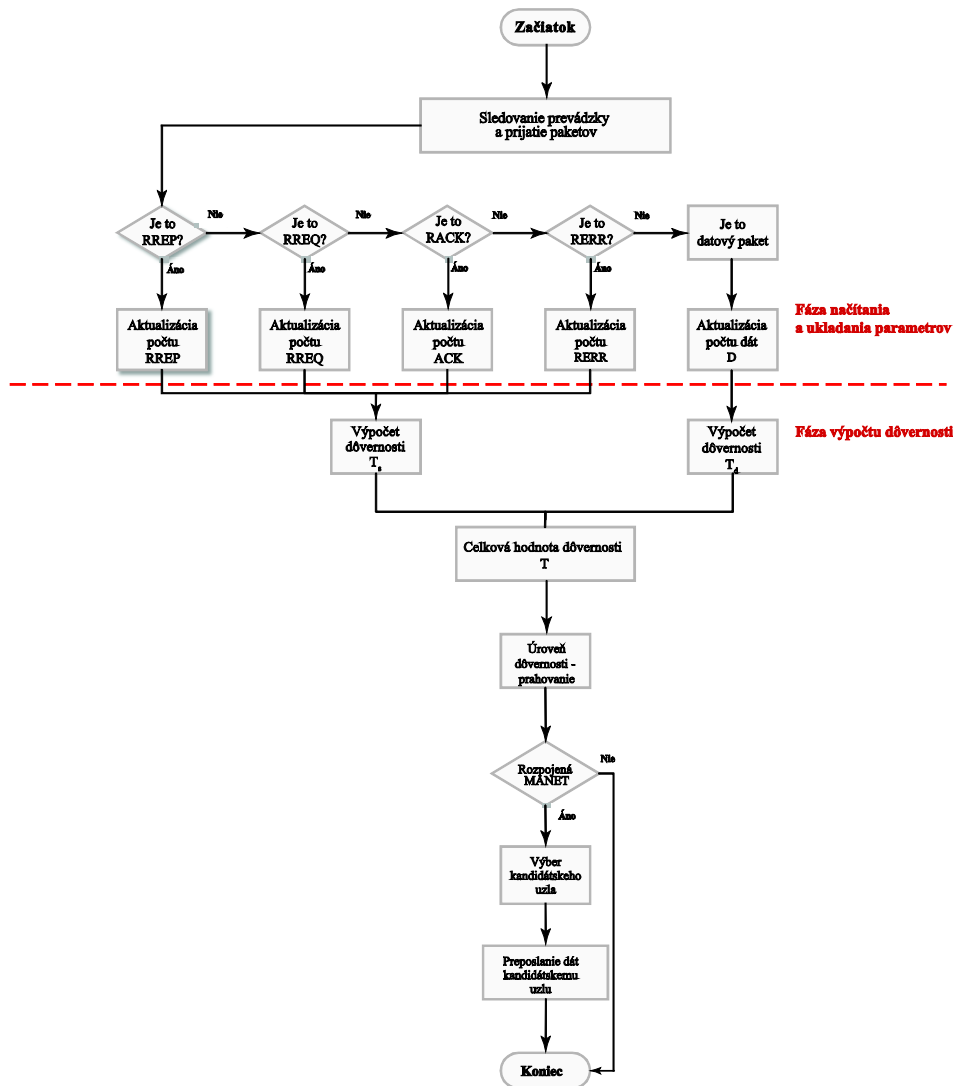
### III. NÁVRH MODELU VÝPOČTU DÔVERNOSTI V HYBRIDNÝCH MANET-DTN

Navrhnutý algoritmus vychádza z priameho modelu pre výpočet dôvernosti. Tento model je založený na predpoklade, že každý uzol v sieti zbiera informácie o ostatných uzloch v sieti. Informácie sú na uzle ukladané a neskôr vstupujú do výpočtu výslednej hodnoty dôvernosti. Výpočet dôvernosti by nemal byť zdĺhavý a mal by mať čo najnižšie nároky pre pamäť a procesor, keďže ide o mobilné zariadenia, ktoré disponujú obmedzenými zdrojmi. Algoritmus je navrhnutý tak, že sa z uzla získavajú informácie o celej prevádzke na sieťovej vrstve referenčného ISO OSI modelu. To znamená, že je sledovaný tok dátových ako aj smerovacích paketov na uzle. Na Obr. 2 je zobrazený vývojový diagram navrhovaného algoritmu pre výpočet dôvernosti. Algoritmus môžeme vo všeobecnosti rozdeliť na dve základné časti:

- a) *Fáza načítavania a ukladania parametrov*
- b) *Fáza výpočtu dôvernosti a rozhodovania*

Je potrebné dodať, že tieto fázy sa vykonávajú vždy, ak je prijatý paket na uzol. Je to z toho dôvodu, aby bola výsledná hodnota dôvernosti vždy aktuálna a aby sa dynamicky menila tak, ako sa mení aj topológia siete, čo súvisí s mobilitou uzlov. Výsledná hodnota dôvernosti potom odráža aktuálny stav zo získaných parametrov v sieti.





Obr. 2 Vývojový diagram výpočtu dôvery pre hybridnú MANET - DTN

### A. Fáza načítavania a ukladania parametrov

Algoritmus pre výpočet dôvery sa spúšťa vždy, ak nastáva zmena v získavaných parametroch, to znamená vtedy, ak je na uzle prijatý paket. Paket prichádzajúci z fyzickej a linkovej vrstvy referenčného ISO OSI modelu je postupne rozbalený a na sieťovej vrstve je to IP paket obsahujúci rôzne informácie. Jednou z týchto dôležitých informácií je *DSR option type*. Táto informácia hovorí o aký druh paketu ide [8]. V navrhnutom algoritme sú použité prichádzajúce pakety ako parametre, ktoré vstupujú do výsledného výpočtu dôvery.

Tieto informácie sú počas fázy načítania a ukladania parametrov ukladané v dátovej štruktúre v pamäti na každom uzle zvlášť. Každý uzol obsahuje informácie o uzloch, s ktorými prišiel do kontaktu a s ktorými si vymieňal smerovacie informácie alebo dáta. Každý uzol je v sieti samostatný a vypočíta si výslednú hodnotu dôvery voči ostatným uzlom. Informácie sú ukladané v štruktúre, ktorá má formu tabuľky.

### B. Fáza výpočtu dôvery a rozhodovania

Po načítaní a aktualizácii parametrov v tabuľke parametrov na každom uzle je potrebné vypočítať a aktualizovať hodnotu dôvery. V tejto fáze sa hodnota dôvery vypočíta zo získaných parametrov. Výpočet dôvery vychádza z priameho modelu pre výpočet dôvery. To znamená, že výsledná hodnota sa teda neposiela ďalej ostatným uzlom ako odporúčanie. Služi lokálne na danom uzle a pomáha mu pri rozhodovaní sa. V našom prípade bude slúžiť vypočítaná hodnota dôvery pri rozhodnutí o posielaní pri situácii, keď dôjde k rozpojeniu spojenia v danej sieti a uzol nebude mať záložnú cestu do cieľového uzla. Samotný výpočet dôvery môžeme rozdeliť do dvoch častí:

- a) Výpočet dôvernosti zo smerovacích paketov  
 b) Výpočet dôvernosti z dátových paketov

Z týchto dvoch čiastkových výpočtov potom dostaneme celkovú hodnotu dôvernosti, ktorá je zapísaná v dátovej štruktúre na uzle. Výpočty jednotlivých čiastkových hodnôt sú uvedené v rovniciach (1) a (2):

$$T_s = W_{rreq} * \left(\frac{R_{req}}{R_{reqc}}\right) + W_{rrep} * \left(\frac{R_{rep}}{R_{repc}}\right) + W_{rerr} * \left(\frac{R_{rerr}}{R_{rerrc}}\right) + W_{rack} * \left(\frac{R_{rack}}{R_{rackc}}\right) \quad (1)$$

$$T_d = W_d * \left(\frac{D}{D_c}\right) \quad (2)$$

kde  $T_s$  a  $T_d$  sú spomínané čiastkové hodnoty dôvernosti. Ďalej  $W_{rreq}$ ,  $W_{rrep}$ ,  $W_{rerr}$ ,  $W_{rack}$ ,  $W_d$  sú konštanty, ktoré slúžia pre normovanie hodnôt dôverností, aby bola potom výsledná hodnota vo zvolenom rozsahu. Konštanty je možné meniť na základe typov sietí a na základe toho, aké váhy chceme pripísať jednotlivým parametrom vstupujúcim do výpočtu dôvernosti. Hodnoty v číateľoch daných rovníc ( $R_{REQ}, R_{REP}, R_{ERR}, R_{ACK}, D$ ) predstavujú počty jednotlivých druhov paketov od konkrétnych uzlov v sieti.  $R_{REQC}, R_{REPC}, R_{ERRC}, R_{ACKC}, D_c$  predstavujú celkové počty paketov prijatých na uzle od všetkých uzlov. Hodnota celkovej dôvernosti predstavuje súčet čiastkových hodnôt ako je to uvedené v rovnici (3):

$$T = T_s + T_d \quad (3)$$

kde  $T$  je výsledná hodnota dôvernosti a  $T_s$  a  $T_d$  sú čiastkové hodnoty dôvernosti uvedené vyššie.

#### IV. SIMULÁCIE A VÝSLEDKY

Simulácie sú prevádzané v programe OPNET modeler [8]. Každá simulácia je opakovaná v 10-tich behoch. Z týchto 10-tich opakovaní sú urobené priemerné hodnoty sledovaných parametrov. Každé opakovanie simulácie je nastavené na trvanie 1 hodinu. Simulácie sú vykonávané pre počty uzlov 20, 40, 60, 80 a 100. Pre každý počet uzlov je menená rýchlosť pohybu mobilných staníc. Dosah jednotlivých mobilných uzlov je 250m. Ako už bolo uvedené hodnoty rýchlostí sú 2, 4, 6 m/s. Sledované boli tieto parametre:

- Priemerný počet preskokov na smerovacej ceste (*Number of hops per route*)
- Prijaté smerovacie informácie (*Total routing traffic received*)
- Odoslané smerovacie informácie (*Total routing traffic send*)

Počas simulácií boli analyzované modely DSR a DSR\_T. Model DSR predstavuje referenčný model smerovacieho protokolu DSR pre MANET [9]. Model DSR\_T je model s implementovaným algoritmom výberu kandidátskych uzlov na báze dôvernosti.

Počas experimentu bol dôraz kladený na simulovanie rozpojení existujúcich spojení medzi mobilnými terminálmi. Tieto rozpojenja boli simulované krátkym vysielacím dosahom mobilných uzlov ako aj náhodným modelom pohybu. Cieľom simulácie je zistenie, ako bude ovplyvňovať navrhnutý algoritmus analyzované parametre.

Tabuľka 1  
 Priemerný počet preskokov v závislosti na rýchlosti mobilných uzlov

Parameter	Počet uzlov / použitý model									
	20 uzlov		40 uzlov		60 uzlov		80 uzlov		100 uzlov	
	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T
2 m/s	3,37	3,03	3,56	3,45	3,46	4,07	4,22	4,28	4,72	4,41
4m/s	3,29	2,33	3,66	3,43	3,69	3,9	4,42	4,321	4,775	4,38
6 m/s	3,49	3,16	3,53	3,61	3,68	3,88	4,27	4,26	4,31	4,95

Z tabuľky 1 vyplýva, že navrhovaný algoritmus nemá vplyv na priemerný počet preskokov. Počet preskokov pri rôznych rýchlostiach je približne rovnaký ako pri nemoifikovanom protokole DSR.

Tabuľka 2  
Prijaté smerovacie informácie v závislosti na rýchlosti mobilných uzlov [bit/s]

Parameter	Počet uzlov / použitý model									
	20 uzlov		40 uzlov		60 uzlov		80 uzlov		100 uzlov	
	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T
2 m/s	1098	1306	2 067	6847	4999	12019	8316,5	14496	16014	28094
4m/s	1567,5	4135,8	3086	2851,2	7243	8843,5	10056	13421	16237	23258
6 m/s	1663	1583,6	2185	9478	6594	9481,5	7897	12329	16104	32103

Tabuľka 3  
Odoslané smerovacie informácie v závislosti na rýchlosti mobilných uzlov [bit/s]

Parameter	Počet uzlov / použitý model									
	20 uzlov		40 uzlov		60 uzlov		80 uzlov		100 uzlov	
	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T	DSR	DSR_T
2 m/s	708	700	1039	1989	2061	2924	4090	5982	8306	9122
4m/s	752	826	2003	2399	3298,5	2160	5931	6561	8784	11394
6 m/s	650	947	1185	3204	2439,5	4957	3910	7625	10000	12694

Získané hodnoty pre parameter priemerný počet prijatých a odoslaných paketov sú znázornené v Tabuľke 2 a 3. V oboch prípadoch je priemerný počet smerovacích informácií väčší. Je to spôsobené tým, že algoritmus je navrhovaný tak, aby pri výpadku spojenia alebo pri strate smerovacej cesty preposlal informácie do naj dôvernejšieho uzla. To však neznamená, že tento uzol je cieľový, alebo že je najbližšie k cieľu. Následne je potrebné posielanie viacerých smerovacích informácií pre opätovné nájdenie cesty od naj dôvernejšieho uzla k cieľovému uzlu v sieti.

## V. ZÁVER

V príspevku je predstavený algoritmus na výpočet dôvernosti susedných uzlov v hybridnej MANET – DTN. Tento algoritmus je následne využívaný v procese výberu kandidátskeho uzla na prenos dát medzi rozpojenými komunikačnými cestami. Algoritmus je implementovaný do smerovacieho protokolu DSR v simulačnom programe OPNET modeler.

Navrhovaný algoritmus operuje konkrétne na sieťovej vrstve ISO OSI modelu. Z výsledkov simulácií vyplýva, že je potrebný väčší čas na spracovanie prichádzajúcich paketov z dôvodu vyhľadávania a aktualizácie štruktúry parametrov a z dôvodu výpočtu dôvernosti. Na druhej strane algoritmus neovplyvňuje počet odoslaných a prijatých dátových paketov, a tým nezaťažuje sieť nadbytočnými informáciami.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## REFERENCIE

- [1] S. Corson, J. Macker, "Mobile Ad hoc Networking (MANET)", RFC, 2501, 1999, 12 s.
- [2] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, H. Weiss, "Delay-Tolerant Networking Architecture", RFC 4838, 2007, 36 s
- [3] Ch. Raffelsberger, H. Hellwagner, "Overview of Hybrid MANET-DTN Networking and its Potential for Emergency Response. Operations Electronic Communications of the EASST," vol. 56, pp.1-12, 2013.
- [4] K. Govindan, P. Mohapatra, "Trust Computations and Trust Dynamics in Mobile Adhoc Networks: A Survey. Communications Surveys and Tutorials", IEEE, Vol. 14, No. 2, 2012, pp. 279 - 298, ISSN: 1553- 877X
- [5] A. Pirzada, C. McDonald, "Trust establishment in pure ad-hoc networks. Wireless Personal Communications, No. 37, Vol.1-2, 2006, pp. 139–168.
- [6] S. Buchegger, J.L. Boudec, "A robust reputation system for P2P and mobile ad-hoc networks". 2nd Workshop on Economics of Peer-to-Peer Systems, 2004.
- [7] J.H. Cho, A. Swami, I.R. Chen, "A survey of trust management in mobile ad hoc networks". IEEE Commun. Surveys Tuts., 2011.
- [8] Z. Lu, H. Yang: Unlocking the Power of OPNET modeler. Cambridge University Press, 2012, 238 s. ISBN 9781139502054
- [9] D. Johnson, Y. Hu, D. Maltz, "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4", RFC4728. IETF, 2007, 107 s.

# PRoPHET protocol for DTN in OPNET modeler

<sup>1</sup>Ján Papaj, Roman Palitefka

<sup>1</sup> Department of Electronics and Multimedia Communications, Technical University of Košice, Faculty of Electrical Engineering and Informatics, Letná 9, 042 00 Košice, Slovak Republic

<sup>1</sup>jan.papaj@tuke.sk

**Abstract** — Delay Tolerant Networks (DTN) are ad-hoc networks in which end-to-end connectivity between source and destination may never be present. In DTNs, the store-carry-and-forward concept is used for solving the problem of intermittent connectivity. Behavior of such networks is verified by real models, computer simulation or combination of the both approaches. Computer simulation has become the primary and cost effective tool for evaluating the performance of the DTNs. OPNET modeler is our target simulation tool and we wanted to spread OPNET's simulation opportunity towards DTN. We implemented bundle protocol to OPNET modeler allowing simulate cases based on bundle concept as epidemic forwarding which relies on flooding the network with messages and the forwarding algorithm based on the history of past encounters (PRoPHET). The implementation details will be provided in article.

**Keywords** — DTN, MANET, PRoPHET, OPNET modeler.

## I. INTRODUCTION

DTN (Delay Tolerant Network) are one of the most effective communication methods for disasters. DTN is the “*challenged computer network*” approach that is originally designed from the Interplanetary Internet. The data transmission is based upon the store - carry- forward model. The protocol carrying data packets under a poor network environment such as space. If mobile nodes such as smart phones or wireless cars are used for carrying data by the DTN method, it is supposed that significant disaster information would be able to transfer to inside and outside the disaster areas [1].

Existing DTN routing protocols evolved from enabling the transfer of any amount of data to carefully selecting intermediate nodes to efficiently carry information. Forwarding schemes were adapted over time to address different performance measures: delivery ratio, message latency, and overhead. The design of DTN routing algorithms may be application-specific, but generally all schemes should balance the overhead from redundant copies with successful delivery and minimal delay. In this work we emphasize the need for multi-objective optimization to better understand performance tradeoffs in opportunistic networks (Fig. 1).

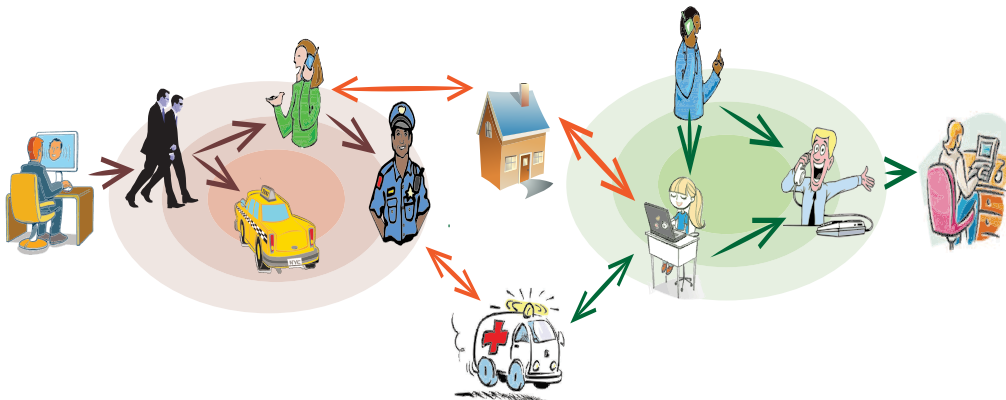


Fig. 1 Delay Tolerant Network

Open source simulation tools such as OMNET++, NS2, NS3 and ONE seemed to be mature for packet - based communication and they include MANET support, bundle architecture and mobility models [3]. On the other hand, proprietary tools as OPNET if we not take into account relatively good MANET support, does not include any compact package for DTN simulation. Benefits of the OPNET is simplicity of the creation network topology, existence of various packages and API interface allowing easy shaping OPNET to various network requirements. New emerging applications and networks as DTNs are receiving increasing attentions from both academia and industry. Demand for new verification tools is increasing [2].

OPNET supports Discrete Event Simulation (DES) and hybrid simulation modes (combination of the mathematical analysis and explicit simulation). The use of the DES simulation provides more accurate and realistic view on object of the interest [3]. DES is typical method used for large-scale simulation studies with extremely detailed model but on the other hand with significant requirements for computing power [6].

Hence OPNET modeler is our target simulation tool we developed bundle protocol extension and the forwarding algorithm based on the history of past encounters PROPHET (Probabilistic Routing Protocol using History of Encounters) for OPNET allowing simulate multiple network scenario based on bundle data unit. The paper is organized as follows. First part „Simulation Model“ is introduction to our model, we continue with the key parts of the model „Bundle protocol” and „PROPHET forwarding algorithm”. In „Conclusions” we simple summarize our work and paper ends with „References“.

## II. PROPHET NODE MODEL IN OPNET MODELER

Overall architecture of our model is shown on Figure 2. A model of the device such as mobile terminal, wireless station, MANET station, and so on is in OPNET referred as “Node Model”. It is modified version of the MANET’s node model `manet_station_adv`. The key parts of the model are blocks `bundle` and `hello_dispatcher`, remaining parts, as can be deduced from its names: `ip_encap`, `ip`, `arp`, `wireless_lan_mac`, `wlan_port_rx` and `wlan_port_tx` represent blocks of the lower layers of the OSI model. We skip the description of these blocks and interested readers we refer to OPNET document [3].

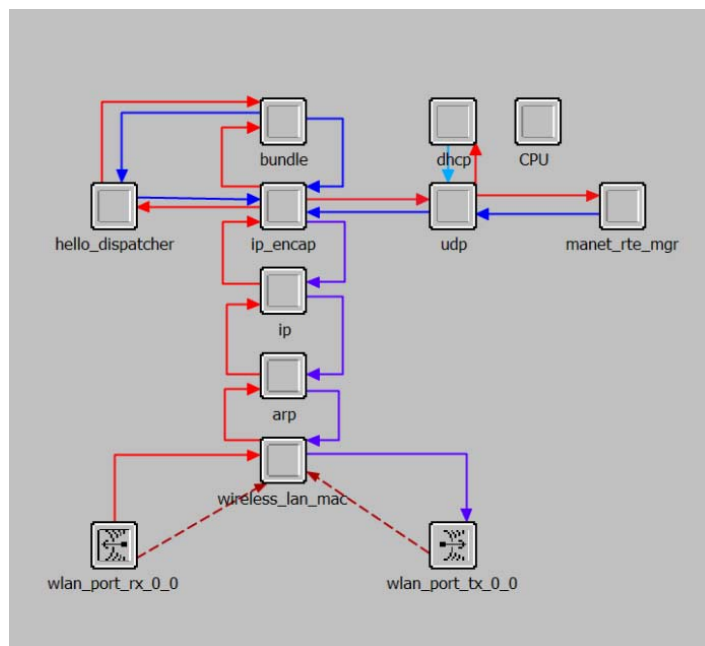


Fig. 2 Implementation of the PROPHET to the simulator OPNET modeler – node model

### A. Bundle layer for PROPHET in OPNET modeler

Bundle protocol describes the end-to-end protocol for the exchange of messages (bundles), end-to-end addressing and acknowledgments [4]. The bundle is data unit created by application which contains both data and meta-data.

Our implementation of the bundle protocol is represented by the block „*bundle*“. It contains in the standard bundle functions also application layer. The application layer is represented by the configurable packet generator. There is possible to configure the start time of the packet generation, the inter-arrival time of the packets, packet size and destination address. The destination address is specified directly or random algorithm can be used. All requirements on bundle protocol can be found in [5].

If the application layer has new bundle to transmit, it is immediately stored to „*bundle buffer*“ with additional information as bundle creation time, expiration time, priority and sequence number. The sequence number is used in ACK message by the destination to confirm successful reception of the bundle. The bundle creation time and expiration time are used by the „*bundle buffer management*“ to maintain bundles in buffer. Confirmed bundles, expired bundles or bundles over predefined time threshold are removed from bundle buffer [6].

New bundle immediately triggers a connection information request to „*Hello dispatcher*“ with destination address included. If connection exists the list of the nodes with delivery predictability are returned by „*Hello dispatcher*“. The block „*bundle*“ is now starting to transmit all stored bundles destined for destination address (including also transition bundles) to one node with highest delivery predictability or multiple nodes with the best delivery predictability values. The number of the destinations selected to send bundle is configuration parameter [5].

#### B. Forwarding algorithm for PRoPHET in OPNET modeler

In OPNET modeler, the „*Hello dispatcher*“ block is our implementation of the PROPHEET forwarding protocol. „*Hello dispatcher*“ is responsible for initiation of the communication between two node. In specified time interval broadcasts „*hello*“ packet to whole network. The „*hello*“ packet is the first communication packet and is followed by the delivery predictability table exchange between nodes and the local transition probabilities calculation as specified in [3].

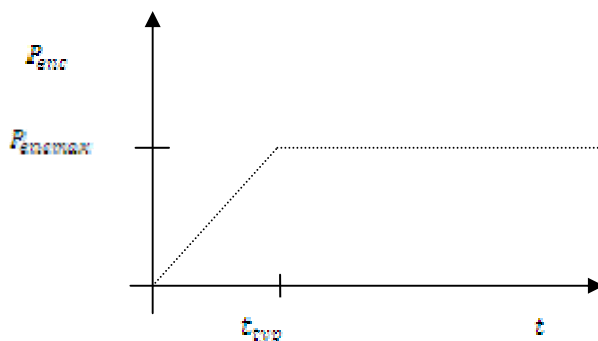


Fig. 3 Graf of the delivery predictabilities for PRoPHET

PRoPHET is specifically intended to provide routing services in a bundle network environment. PRoPHET is mentioned to run over reliable TCP, but if bundles are provided with sequence numbers, together with a capability for positive acknowledgements it allows PRoPHET to operate over an unreliable protocol such as UDP or directly over IP. Implementation of the protocol directly over IP decreases the implementation overhead, so we decided to operate over IP with bundle acknowledgment sent on the successful bundle reception by the destination.

$$P(A,B) = P_{old}(A,B) + (1 - \delta - P_{old}(A,B)) * P_{enc} \quad (1)$$

When two nodes A and B meet each other, they update the delivery predictability for each other. Nodes that are often in encounter have a high delivery predictability. Based on the values of the delivery predictability is performed decision to forward or not forward bundle to a certain node. The PRoPHET operation we can shortly describe by these equations. On encounter node A calculates for each known destination B delivery predictability  $P(A,B)$  (1).

Where

- $0 \leq P(A,B) \leq 1$ ,
- $P_{old}(A,B)$  is the value that was stored before the encounter,
- $\delta$  is a small positive number that effectively sets an upper bound for  $P(A,B)$  in a

way that predictabilities between different nodes stay strictly less than 1,

- $P_{enc}$  is a function of the interval since the last encounter resulted in an update of the delivery predictabilities (Fig. 3). If the intervals between encounters  $P_{enc}$  is longer than the "typical" time  $t_{typ}$ ,  $P_{enc}$  is set to a fixed value  $P_{encmax}$ .

The node transmitting its delivery predictabilities needs before the transmission to age delivery predictabilities values to reflect the time spent from last encounter with any other node according to equation (2).

$$P(A,B) = P_{old}(A,B) * \gamma^K \tag{2}$$

$$P(A,C) = \text{MAX}(P(A,C)_{old}, P(A,B) * P(B,C)_{recv} * \beta) \tag{3}$$

Where

- $\gamma$  is aging constant  $0 \leq \gamma \leq 1$
- $K$  is the number of time units elapsed since the last metric aging
- $\beta$  is scaling constant  $0 \leq \beta \leq 1$ , controlling the impact of the transitivity on the delivery predictability
- $P(B,C)_{recv}$  is the value received from the node B

If node A frequently encounters node B but not frequently node C and on other hand B frequently encounters C but not A and A wants to transmit bundle to node C then node B seems to be good transitive node for A. The last equation (3) represents this transitive property:

Parameter  $\delta$ ,  $P_{encmax}$ ,  $\gamma$ ,  $\beta$  are configurable depending on the usage scenario, in [4] authors provide some recommended default values.

### III. SIMULATIONS AND RESULTS

We ran our experiments using the OPNET modeler network simulator [7]. OPNET Modeler enables to simulate mobile ad-hoc networks with different routing protocol. In our case, the Dynamic Source Routing protocol was used as a reference value [8],[9]. The main objective of this test is to simulate and analyze the functionalities of the routing protocol DSR for MANET and PRoPHET designed for the DTN in the case when the communication paths are disconnected for short time interval. It is interval when DSR routing protocol for MANET cannot find a communication path with another mobile nodes.

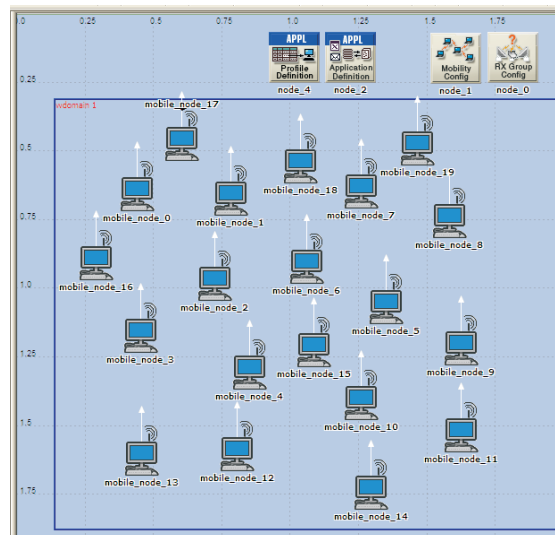


Fig. 4 Example of the simulation scenario for 20 mobile nodes in OPNET modeler

#### A. OPNET modeler simulations

Simulated parameters were used to establish the average delay, average number of routing data and average number of hops. The number of nodes is varying and considered 20, 40, 60, 80 and 100 mobile nodes. The all simulation parameters are considered as in the Table 1. The following simulation scenarios were used during the simulations, namely (Fig. 4):

- *MANET model* – MANET network with implemented DSR protocol without the possibility to find communication paths between mobile nodes.
- *DTN model* – model with implemented PRoPHET protocol and protocol enable to find a communication path if the links are disconnected during the transportation of the packets.

Table 1  
 Simulation setup parameters for OPNET modeler

Parameters	Values
Routing protocols	DSR, PRoPHET
Simulation area	2000 m x 2000 m
Time of simulation	3600 s
Transmission range	250 m
Number of nodes	20, 40, 60, 80, 100
Speed of nodes	0 -10 m/s
Model placement	Default
Mobility model	Default Random Waypoint with pause time10 s
Number of values per simulation	1000

The average delay is assumed to be the average amount of the time that is necessary for transmission of the packet between source and destination nodes. Average number of hops represents necessary to find end-to-end connection between source and destination mobile node. The average number of routing data gives us information about how many data the routing protocol needs to send to the network until the destination node will find.

#### B. OPNET modeler simulation results

In simulations, we focus on the analysis of these average values. The main idea of the experiment is knowledge how the protocol PRoPHET for DTN will affect the functionality of the network in the case of the temporary disconnected MANET network. The disconnection of the communication paths is simulated by a short transmission range of the mobile nodes and random mobility waypoint model with random speed of the mobile nodes within the range 0 - 10 m/s.

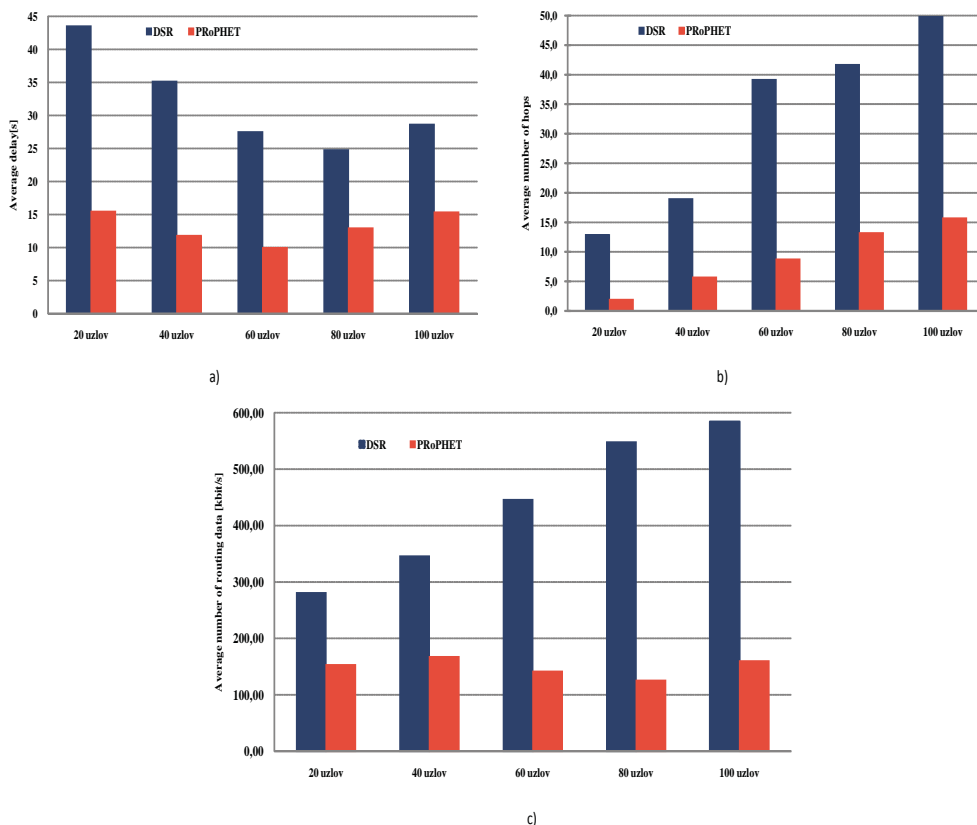




Fig.5 Simulation results: a) The average delay for MANET and DTN, b) The average number of the hops for MANET and DTN, c) The average number of routing data for MANET and DTN

The delays experienced by the protocols are a crucial factor contributing to the performance of the protocols. We can see in Fig. 5a that the DSR protocol for MANET has the average delay much higher than the PROPHET designed for DTN network. It is clear from the Fig. 5a that the average delay decreased with the number of nodes. The delay was lowest for PROPHET protocol designed for DTN. It is because the DSR routing protocol always sends the routing protocols in order to find communication paths and on the other side PROPHET protocol unicast only the packet for the nodes which are directly connected to the mobile node. The values of average delay for MANET are almost 2 times higher than values for DTN.

The second parameter analyzed is average number of hops. We can see that the DSR protocol provides the routing mechanism if the communication paths are disconnected which is resulting in a significant number of the hops for networks with the highest number of the mobile nodes (Fig. 5b). The PROPHET protocol on the other side enables to find destination node with lower number of the hops. A protocol for DTN shows the best performance as MANET when the communication paths are disconnected (Fig. 5b).

The third one analyzed parameter is the average number of routing data. Fig. 5c shows the results measured in kbit/s versus different number of the mobile nodes. In all cases the protocol for DTN network sends the lower number of the data. In the case of disconnected communication paths, the MANET network with a DSR routing protocol to send a lot of routing packet to the network in order to find connection to the other mobile nodes and for this reason the average number of the routing data is highest for all simulation scenarios in comparison with DTN. Forwarding algorithms for DTN are based on directly transportation of the data if the mobile node is the transitions range of the source mobile node.

#### IV. CONCLUSIONS

The article gives us the basic performance analysis of the MANET and DTN networks in an OPNET modeler simulator environment. Based on collected results we can conclude the forwarding algorithm for DTN are optional way how the data could be transported in a disconnected environment where the routing protocol for MANET have no mechanism to transport of the data. On next research we will focus on the methods how to combine routing properties of the MANET and DTN in order to increase the performance of the network.

New emerging applications and networks are receiving increasing attention from both academia and industry. The demand for new verification tools is increasing. For OPNET modeler as our target simulation we have developed the bundle protocol extension and the forwarding algorithm based on the history of past encounters PROPHET. The implementation introduction together with bundle protocol and PROPHET forwarding algorithm description are provided.

#### ACKNOWLEDGMENT

This work is the result of the project implementation: Development of the Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120030) supported by the Research & Development Operational Program funded by the ERDF.

#### REFERENCES

- [1] M.K. Denko, " *Mobile Opportunistic Networks*", Taylor and Francis Group, LLC, ISBN: 13: 978-1-4200-8813-7, 2011
- [2] Y. Cheng, D.Y. Eun, Z. Qin, M. Song, K. Xing, " *Wireless Algorithms, Systems, and Applications*", 6<sup>th</sup> International Conference, WASA 2011, Springer-Verlag Berlin Heidelberg 2011, ISBN 978-3-642-213489-3, 2011
- [3] Z. Lu, H. Yang, " *Unlocking the Power of OPNET Modeler*", Cambridge University Press 2012, ISBN 978-0-521-19874-5 Hardback, 2012
- [4] A. Lindgren, A. Doria, E. Davies, S. Grasic, S., " *Probabilistic Routing Protocol for Intermittently Connected Networks*", ISSN: 2070-1721, <http://tools.ietf.org/html/rfc6693>, August 2012
- [5] K. Scott, S. Burleigh, " *Bundle Protocol Specification*", <http://tools.ietf.org/html/rfc5050>, November 2007
- [6] P. Hui, J. Crowcroft, E. Yoneki, "BUBBLE Rap: Social-Based Forwarding in Delay-Tolerant Networks", *Mobile Computing, IEEE Transactions on*, pp. 1576-1589, Vol. 10, issue 11, 2011.
- [7] OPNET Modeler Simulation Software, <http://www.opnet.com>.
- [8] A. Čížmár, L. Doboš, J. Papaj, " *Security and QoS Integration Model for MANETs*", In *Computing and Informatics* Vol. 31, No.5, pp. 1025-1044, ISSN: 1335-9150, 2012.
- [9] J. Papaj, L. Doboš, A. Čížmár, " *Routing Strategies in Opportunistic Networks*", *Journal of Electrical and Electronics Engineering*, Vol. 5, No. 1, pp. 167-172, ISSN 1844-6035, 2012.

# Riešenie elektromagnetických polí s využitím ANSYS

<sup>1</sup>Alexander MÉSZÁROS, <sup>2</sup>Marek MAJERSKÝ, <sup>3</sup>Ján ZBOJOVSKÝ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>alexander.meszaros@tuke.sk, <sup>2</sup>marek.majersky@student.tuke.sk,  
<sup>3</sup>jan.zbojovsky@tuke.sk

**Abstrakt** — Práca analyzuje možnosti a výhody numerických metód riešenia elektromagnetických polí, detailnejšie je opísaná metóda konečných prvkov. Je uvedený taktiež stručný opis vybraných softvérových prostriedkov riešenia, s hlavným dôrazom na ANSYS a jeho platformu Workbench, ktorý sa využil na riešenie ukázkových modelových prípadov magnetických polí vo vybraných elektrických zariadeniach a ich okolí.

**Kľúčové slová** — elektromagnetické pole, metóda konečných prvkov, ANSYS, Workbench

## I. ÚVOD

Elektromagnetické polia sa nachádzajú všade okolo nás a vyhnúť sa ich pôsobeniu je v súčasnej dobe nemožné, pričom postupom času počet ich zdrojov sa bude ďalej zvyšuje. Preto je nevyhnutné venovať sa ich riešeniu, skúmaniu a analýze. Za zdroj elektromagnetických polí je možné považovať v podstate všetko elektricky vodivé, čím prechádza elektrický prúd ako napríklad elektrické vedenia, transformátory, rôzne elektrické spotrebiče, či elektrické inštalácie v bytoch. Každý deň sme vystavovaní pôsobeniu elektromagnetických polí z týchto zdrojov, ktoré ich generujú a všetky tieto faktory môžu negatívne vplývať nielen na ľudský organizmus ale aj na rôzne iné elektrické zariadenia. Avšak využitie elektromagnetických polí má aj kladnú stránku. V súčasnosti sa stretávame s využívaním elektromagnetizmu takmer v každej oblasti spoločenského života, napr. v telekomunikáciách, zdravotníctve, výskume a pod., a jeho využitie má stále väčší význam.

Táto práca sa venuje opisu niektorých matematických (numerických) metód a tiež prostriedkov, ktoré slúžia na ich riešenie. Boli vytvorené matematické modely v programe ANSYS a jeho platforme Workbench, v ktorom boli namodelované rôzne prípady šírenia elektromagnetických polí vo vybraných elektrických zariadeniach a ich okolí.

## II. METÓDY RIEŠENIA ELEKTROMAGNETICKÝCH POLÍ

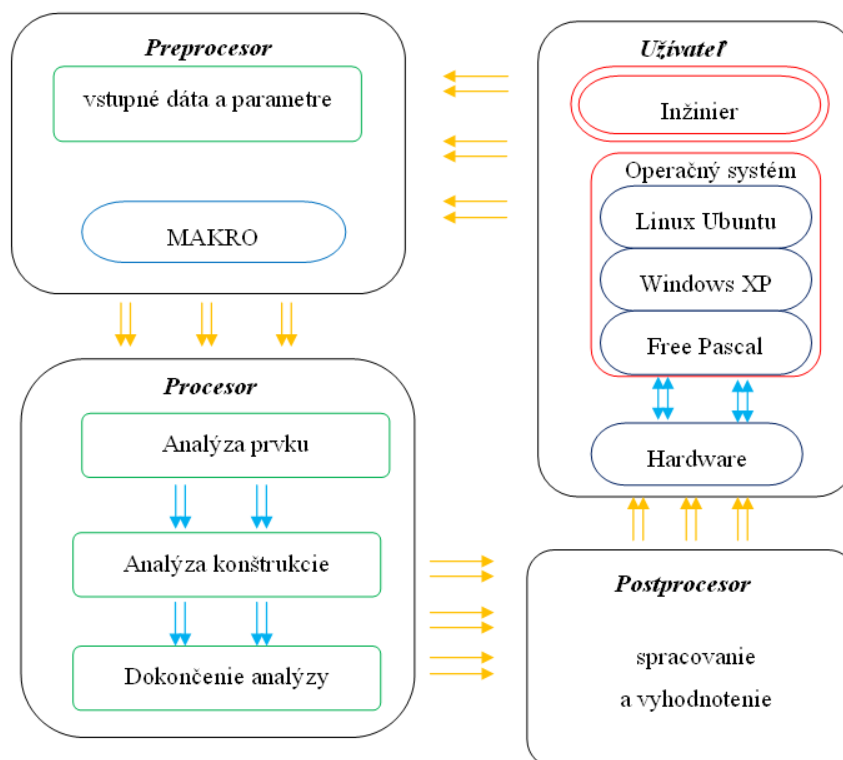
Pri riešení úloh elektromagnetických polí sa využívajú rôzne metódy a prostriedky. Vo všeobecnosti je možné rozdeliť metódy na analytické a numerické. Pri riešení jednoduchých úloh poľa sú využívané analytické metódy, naopak pri zložitejších úlohách je potrebné riešiť úlohy numericky. Vzhľadom na náročnosť riešení pri niektorých úlohách, je pri výpočtoch numerických metód využívaná výpočtová technika spolu so špeciálnymi aplikáciami na to určenými, prípadne aj programovacími jazykmi.

Možnosti analytického riešenia polí sú výhodné pre relatívne široké spektrum čiastkových problémov, avšak zlyhávajú pri riešení reálnych úloh. Hlavnou príčinou je zložitosť geometrie problému, nelineárne charakteristiky materiálov, potreba riešenia úloh v priestore (3-D) alebo pridanie ďalšieho rozmeru – času do úlohy [1]. Práve tu nastupujú riešenia numerické.

Hlavnou podstatou numerických metód je to, že transformujú systém diferenciálnych rovníc na systém algebrických rovníc. Diferenciálne rovnice sú dôležité pri definovaní neznámych funkcií poľa na konečnej ohraničenej oblasti. Riešením diferenciálnej rovnice sa nájde neznáma funkcia poľa, ktorá musí spĺňať počiatočné a okrajové podmienky. Hlavnou podstatou väčšiny numerických metód riešenia diferenciálnych rovníc je, že sa pre neznámu funkciu poľa vhodne zvolí náhradná funkcia vzt'azných súradníc, ktorá pozostáva z určitého typu funkcie (polynómy, goniometrické funkcie) a voľných koeficientov. Jednotlivé numerické metódy sa

navzájom líšia náhradnou funkciou, fyzikálnym významom voľných koeficientov a spôsobom ich určenia. Fyzikálny model, ktorý je daný geometriou a zdrojmi poľa, sa transformuje na matematický model, čomu zodpovedá popis problému rovnicami (napr. Maxwellove). Aby bol matematický model jednoznačne definovaný pri hľadaných veličinách na jeho vonkajšom ohraničení, je potrebné určiť okrajové podmienky úlohy. Potom sa volia strategické body (diskretizácia) a tým vzniká numerický model. Pri riešení numerického modelu sa pracuje s numerickými metódami alebo je možné použiť vhodný výpočtový program. V algebrickej realizácii úlohy sa formulujú systavy rovníc, ktoré popisujú problém do formy matic. Nasleduje počítačové riešenie úlohy, výpočet veličín a spracovanie výsledkov [1], [2].

Medzi najznámejšie numerické metódy patria: metóda konečných diferencií, metóda hraničných prvkov, momentová metóda a metóda konečných prvkov (MKP). Prudký rozvoj MKP začal počiatkom 60-tych rokov 20. storočia a jej vývoj nie je ukončený ani v súčasnej dobe. Najskôr sa táto metóda rozvíjala len pre oblasť mechaniky pevných a poddajných telies. Postupom času sa ukázalo, že je ju možné aplikovať aj na riešenie iných úloh ako je teplotné pole, elektrina a magnetizmus a pod. V súčasnosti patrí medzi najuniverzálnejšie, najrozšírenejšie a najefektívnejšie numerické metódy riešenia úloh poľa [2]. MKP transformuje výpočet primárnej neznámej v teoreticky nekonečnom počte bodov oblasti, na výpočet jej hodnôt v konečnom počte bodov. Práve tento postup ju odlišuje od iných numerických metód, pretože vhodne zvolené náhradné funkcie sa volia len pre oblasť konečného rozmeru (konečný prvok). Dôležité je to, že zvolené náhradné funkcie sú pre všetky prvky danej skupiny rovnaké. Každý konečný prvok je medzi sebou prepojený uzlom a tým vytvárajú určitú sieť. Vzhľadom na to, že pri riešení niektorých zložitejších úloh by bol numerický postup riešenia metódou konečných prvkov veľmi komplikovaný, využíva sa spolupráca s výpočtovou technikou na základe algoritmu podľa obr.1. [2], [3], [4].



Obr. 1 Algoritmus výpočtu metódou konečných prvkov

### III. PROSTRIEDKY RIEŠENIA ELEKTROMAGNETICKÝCH POLÍ

V súčasnosti existuje mnoho prostriedkov na riešenie elektromagnetických polí, ktoré pracujú na báze numerických metód, čo uľahčuje riešenie a skracaie čas. Pre modelovanie elektromagnetických polí je možné použiť niekoľko programových prostriedkov. Medzi najznámejšie patria ANSYS, FEMM, MATLAB a Comsol Multiphysic.

#### A. FEMM (Finite Element Method Magnetics)

Je jednoduchý a nenáročný programový balík pre riešenie 2-D rovinných a osovo symetrických lineárnych a nelineárnych, časovo harmonických úloh nízko-frekvenčných elektromagnetických polí, lineárnych elektrostatických polí a úloh v ustálenom stave tepelných

tokov. Softvér rieši úlohy metódou konečných prvkov. Postup modelovania úloh je možné v programe FEMM rozdeliť na dve časti. Prvou časťou je preprocessor (definovanie), v ktorej sa vytvorí geometria modelu, priradia sa vlastnosti materiálov a určia sa okrajové podmienky. Výhodou je, že FEMM je kompatibilný s CAD programami, čo umožňuje navrhnuť a následne importovať geometriu modelu vytvorenú v týchto programoch, ktoré sú určené pre technické kreslenie. Druhou časťou je postprocessor (vyhodnotenie), kde je možné zobrazit' výsledné riešenie úlohy. Zobrazenie riešenia je možné vykonať v troch režimoch: režim bodovej hodnoty (jednotlivé body riešenej oblasti), režim obrysu (ľubovoľné kontúry oblasti) a blokový režim (subdoména v oblasti riešenia) [5].

#### B. MATLAB (MATrix LABoratory)

Je vysoko výkonný software určený pre numerické výpočty, vizualizácie a programovanie v ľahko použiteľnom prostredí, kde sú všetky úlohy a riešenia vyjadrené matematickým zápisom. MATLAB je široko používaný v oblastiach aplikovanej matematiky, vzdelávania, výskumu, vysokých škôl a priemyslu.

Základnými údajovými prvkami sú matice a vektory. Práve to robí MATLAB obzvlášť užitočným pre lineárnu algebru. MATLAB taktiež slúži pre riešenie algebrických, diferenciálnych rovníc a numerickej integrácie. Jeho súčasťou sú výkonné grafické nástroje, ktorými je možné prezentovať výsledky výpočtov v podobe 2D alebo 3D obrázkov, grafov alebo animácií. Pri riešení úloh je možné využiť veľké množstvo funkcií pre matice a vektory. Taktiež disponuje základnými numerickými algoritmami, medzi ktoré patria interpolácia, aproximácia metódou najmenších štvorcov, korene polynómov, riešenie nelineárnych rovníc a ich sústav, minimalizácia funkcie viac premenných, riešenie diferenciálnych rovníc a ich sústav [6].

#### C. COMSOL Multiphysics

Prostredie Comsol Multiphysics (do r. 2005 FEMLAB) je určené na riešenie časovo závislých alebo stacionárnych úloh (z oblasti akustiky, tepla, elektromagnetizmu...), ktoré sú popísané parciálnymi diferenciálnymi rovnicami v 1D, 2D alebo 3D sústave. Tento programovací jazyk využíva pri riešení úloh metódu konečných prvkov. Poskytuje sofistikované ale zároveň pohodlné nástroje pre geometrické modelovanie.

Comsol Multiphysics umožňuje spoluprácu s programom MATLAB. Keďže program pracuje na základe metódy konečných prvkov, tak pri riešení úlohy je potrebné definovať typ prostredia, ktoré je popísané parciálnymi diferenciálnymi rovnicami. Tak isto je dôležité zadať okrajové podmienky v danom geometrickom modeli. Súčasťou programu sú taktiež knižnice parciálnych diferenciálnych rovníc, ktoré definujú rôzne aplikačné režimy. Výhodou je, že výberom režimu je možné zadať vlastnosti danej oblasti a okrajové podmienky bez zložitých matematických definícií. Podstatou názvu Multiphysics je, že program dokáže kombinovať niekoľko aplikačných režimov do jedného modelu a tým sa vytvorí multifyzikálna úloha [7].

#### D. ANSYS

ANSYS je multifyzikálny program, ktorý je určený pre riešenie lineárnych a nelineárnych úloh rôznych druhov, ktorými sú: štruktúrna analýza, rázové deje, vedenie tepla, prúdenie, elektromagnetické polia, elektrostatika, akustika, lomová mechanika a kompozity. Umožňuje uskutočňovať nielen kontrolné výpočty, ale na ich základe umožní následne previesť optimalizáciu, a to, ako topologickú, tak i citlivostné analýzy. ANSYS ako program je navrhnutý tak, aby riešil jednotlivé úlohy metódou konečných prvkov [8], [9].

Platforma ANSYS Workbench je prostredie, ktoré združuje široké spektrum pokročilých technológií použitých v simulačnom softvéri od spoločnosti ANSYS. Nový inovatívny pohľad na projekt, ktorý automaticky prepojuje programy užívané vo všetkých krokoch simulácie, sprevádza užívateľa celým procesom, dokonca i pri veľmi zložitých multifyzikálnych analýzach. ANSYS Workbench prináša mimoriadnu produktivitu umožňujúcu vyvíjať produkty s použitím simulačných nástrojov pomocou obojsmerného prepojenia s CAD nástrojmi a výkonným automatizovaným nástrojom pre vytváranie sietí, automatickú aktualizáciu projektov, parametrizovanie úloh a integrovanými optimalizačnými nástrojmi [10].

Všeobecný postup modelovania poľa je v programe ANSYS nasledovný:

- Vytvorenie fyzického prostredia – definícia typov prvkov, konštánt, materiálových vlastností a prostredia,
- Tvorba geometrie – vytvorenie daného modelu,

- Priradenie atribútov – priradenie materiálov danému modelu,
- Tvorba siete – diskretizácia,
- Priradenie okrajových podmienok – jednoznačné definovanie oblasti minimálne v jednom uzle,
- Aplikácia záťaže – prúd, magnetické pole, elektrické pole,
- Riešenie,
- Výsledky.

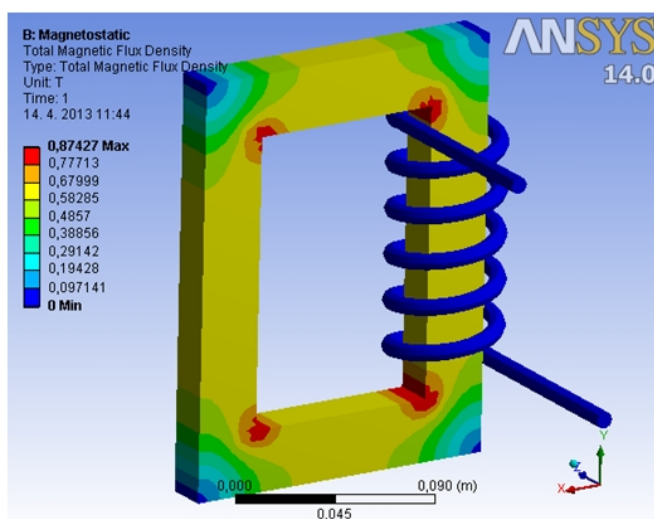
#### IV. MODELOVANIE ELEKTROMAGNETICKÝCH POLÍ VO VYBRANÝCH ELEKTRICKÝCH ZARIADENIACH A V ICH OKOLÍ

Elektrické, magnetické či elektromagnetické polia v praxi môžu byť: stacionárne, kvazistacionárne a nestacionárne.

Stacionárne pole je také pole, ktorého všetky charakteristické veličiny sú časovo nezávislé. Typickým príkladom takého poľa môže byť pole vytvorené elektrickým prúdom, pri ktorom sa náboje pohybujú konštantnou rýchlosťou (jednosmerný prúd). Tokom takého prúdu sa vytvára elektrické prúdové pole a stacionárne magnetické pole.

Ako príklad je navrhnutý model cievky, do ktorej je vložené železné jadro. Cievka je napájaná zdrojom jednosmerného prúdu a tým vytvára vo svojom okolí stacionárne elektromagnetické pole. Geometria modelu je vytvorená v platforme Workbench v 3-D zobrazení. Po vytvorení geometrie je potrebné určiť typ analýzy. V tomto prípade je zvolená analýza typu "Magnetostatic", ktorá spĺňa všetky kritéria pre celý proces modelovania, riešenia a vyhodnotenia výsledkov. Ďalej je potrebné priradiť modelu typy materiálov z knižnice, ktorú program obsahuje. Vytvorenie siete je realizované automatickou metódou, čo znamená, že program si sám vhodne zvolí typy konečných prvkov a hustotu siete (v prípade potreby je možné manuálne nastavenie). Určenie okrajovej podmienky, ktorou sa jednoznačne určí riešená oblasť a jej hranica sa v programe realizuje aplikovaním "Magnetic flux parallel" (nulový potenciál, ktorý predstavuje spôsob ohraničenia poľa plochou vzduchu. Posledným krokom je priradenie záťaže, čo bolo v tomto prípade určené zadaním veľkosti elektrického prúdu. Po tejto príprave nasleduje samotné programové riešenie úlohy a vyhodnotenie výsledkov.

Na obr. 2 je vidieť rozloženie magnetickej indukcie v magnetickom obvode, kde najväčšia deformácia magnetického poľa je v rohoch železného jadra. Pri jednosmernom elektrickom prúde 10 A pretekajúcej cievkou, vznikla v železnom jadre magnetická indukcia s maximálnou hodnotou 0,87 T.

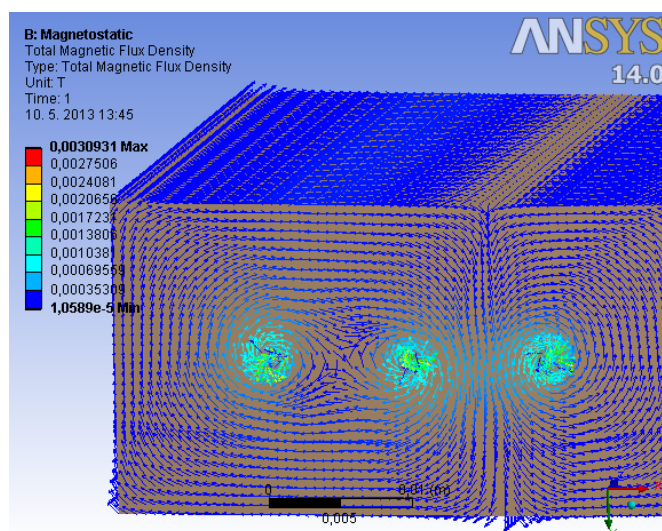


Obr.2 Rozloženie magnetickej indukcie v magnetickom obvode

Kvazistacionárne elektromagnetické pole je pole, v ktorom sa dejú pomalé zmeny smeru pohybu elektrického náboja (nízkofrekvenčný striedavý prúd). Typické kvazistacionárne elektromagnetické pole sa vytvára v elektrických rozvodoch a sieťach pri frekvencii 50Hz resp. 60 Hz.

Na obr. 3 je vektorovo znázornená magnetická indukcia poľa vytvoreného modelu troch vodičov, ktorá určuje tvar a orientáciu magnetického poľa. Ľavým a stredným vodičom preteká elektrický prúd rovnakého smeru (do zobrazenej oblasti), pravým vodičom preteká elektrický prúd v opačnom smere (zo zobrazenej oblasti). Ako je vidieť z obrázku, tak medzi vodičmi,

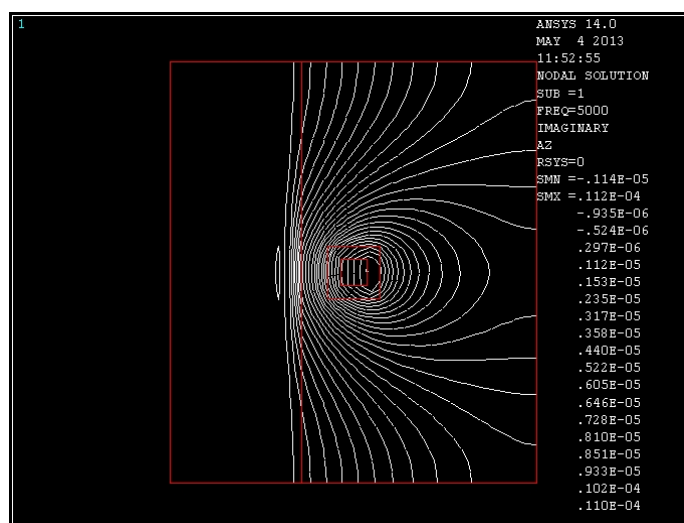
ktorými preteká elektrický prúd rovnakého smeru (ľavý a stredný) je indukcia magnetického poľa menšia, čo znamená, že v týchto sa účinky magnetického poľa zoslabujú (magnetické pole je redšie). V prípade opačného smeru elektrického prúdu pretekajúceho vodičmi (stredného a pravého) je indukcia magnetického poľa medzi vodičmi väčšia, čiže magnetické pole sa v týchto miestach zosilňuje (magnetické pole je hustejšie).



Obr.3 Vektorové znázornenie magnetickej indukcie modelu troch vodičov napájaných trojfázovým harmonickým prúdom

Pre prípad nestacionárneho elektromagnetického poľa je vytvorený model jednovávitovej cievky, ktorou tečie vysokofrekvenčný elektrický prúd. Tento proces je využívaný v praxi pre indukčný ohrev, ktorým sa dajú rýchlo dosiahnuť požadované vlastnosti pri spracovávaní materiálov. Princíp indukčného ohrevu spočíva vo vysokofrekvenčnom prúde, ktorý tečie vo vodiči (tvar cievky, špirály a podobne) okolo ohrievaného materiálu a vplyvom elektromagnetickej indukcie sa z cievky indukujú vírivé prúdy do elektricky vodivého materiálu, ktoré ho ohrievajú. Reguláciou frekvencie je možné meniť to, do akej hĺbky prierezu sa materiál ohrieva. Model je vytvorený v klasickom ANSYS-e v 2-D zobrazení.

Na obr. 4 je siločiarami znázornený tvar magnetického poľa v okolí modelu. Ako je možné vidieť, tak najväčšia hustota siločiar (intenzita) magnetického poľa je práve medzi cievkou (induktorom) a tyčovým polotovarom (ohrievaným telesom), pretože práve v týchto miestach prebiehajú zmeny elektromagnetického poľa.



Obr.4 Znázornenie siločiar magnetického poľa v okolí modelu

## V. ZÁVER

Na báze MKP pracuje niekoľko softvérových prostriedkov, ktoré boli stručne charakterizované a sú veľmi často využívané pri riešení najmä náročnejších úloh, čo



oslobodzuje daného užívateľa od zložitých a časovo náročných výpočtov. Hlavnou úlohou bolo vytvoriť niekoľko matematických modelov v už spomínanom programe ANSYS a jeho platforme Workbench, ktorý bol v tejto práci využitý pre štúdiu rozloženia elektromagnetických polí v elektrických zariadeniach a ich okolí. Prostredie Workbench pôsobí pre užívateľa prehľadnejšie a tvorba zložitejších geometrií modelov (3-D) je v ňom jednoduchšia, oproti klasickému ANSYS-u. Pri daných typoch úloh, ktoré boli riešené v tejto práci sa výsledky jednotlivých analýz nie vždy zhodovali s teoretickými predpokladmi. Tieto problémy boli pravdepodobne spôsobené nedokonalosťou aplikácie alebo staršou verziou, nakoľko je neustále v procese rozvíjania a vylepšovania.

Modelovanie je v klasickom ANSYS-e trochu odlišné ako vo Workbench-i, avšak na tej istej úrovni, čo sa týka náročnosti. Klasický ANSYS je vhodnejší pri riešení geometricky jednoduchších úloh (1-D, 2-D). Výsledky riešených úloh sa zhodovali s teoretickými predpokladmi a nedošlo k žiadnym väčším problémom počas procesu modelovania.

V konečnom porovnaní klasického ANSYS-u a Workbench-u je možné vysloviť záver, že obidve aplikácie majú svoje výhody aj nevýhody, v závislosti od toho aký typ úlohy je potrebné riešiť. Klasický ANSYS pôsobí konzervatívnejšie, naopak Workbench disponuje modernejším prostredím. Každopádne celý programový balík je veľmi vhodným prostriedkom a alternatívou ako riešiť úlohy z oblasti elektromagnetických polí, aj napriek niektorým nedostatkom, ktoré sú ale vývojármí postupne odstraňované.

#### POĎAKOVANIE

Táto práca vznikla vďaka podpore v rámci operačného programu Výskum a vývoj pre projekt: Ochrana obyvateľstva Slovenskej republiky pred účinkami elektromagnetického poľa, s kódom ITMS: 26220220145, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.



Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ.

#### LITERATÚRA

- [1] MAGA, Dušan – HARTĀNSKÝ, René: Numerické riešenia [online]. Dostupné na internete: <[http://www.maiki.ic.cz/2sem4roc/numries/num\\_ries\\_BRNO.pdf](http://www.maiki.ic.cz/2sem4roc/numries/num_ries_BRNO.pdf)>.
- [2] GOGA, Vladimír: Počítačové riešenie polí [online]. Bratislava: STUBA, FEI. Dostupné na internete: <[http://aladin.elf.stuba.sk/Katedry/KMECH/slovakversion/Predmety/PRP/goga/PRP\\_1\\_1\\_uvod%20do%20PRP.pdf](http://aladin.elf.stuba.sk/Katedry/KMECH/slovakversion/Predmety/PRP/goga/PRP_1_1_uvod%20do%20PRP.pdf)>.
- [3] BĪREŠ, Pavol: Numerické metódy výpočtu elektromagnetického pole: Diplomová práca. Brno: VUTBR, Fakulta elektrotechniky a komunikačných technológií, 2010. 65 s. Dostupné na internete: <[http://www.vutbr.cz/www\\_base/zav\\_prace\\_soubor\\_verejne.php?file\\_id=26531](http://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=26531)>.
- [4] KUTIŠ, Vladimír: Základy modelovania a simulácii [online]. Bratislava: STU, FEI, 2006. Dostupné na internete: <[http://www.energetici.tym.sk/system/docasne/upni\\_60.pdf](http://www.energetici.tym.sk/system/docasne/upni_60.pdf)>.
- [5] JIMENEZ, Roberto: FEMM [online]. Dostupné na internete: <<http://femm.software.informer.com/4.2/>>.
- [6] What is MATLAB [online]. Dostupné na internete: <<http://cimss.ssec.wisc.edu/wxwise/class/aos340/spr00/whatismatlab.htm>>.
- [7] COMSOLMultiphysics [online]. Praha: FEL, ČVUT, 2008. Dostupné na internete: <[http://old.elmag.org/lib/exe/fetch.php/wiki:user:cifra:uvod\\_do\\_cm.pdf](http://old.elmag.org/lib/exe/fetch.php/wiki:user:cifra:uvod_do_cm.pdf)>.
- [8] VRBKA, Martin – VAVERKA, Michal: Metóda konečných prvků [online]. Brno: VUTBR, ústav konštruování. Dostupné na internete: <[http://old.uk.fme.vutbr.cz/kestazeni/MKP/prednaska1\\_mkp.pdf](http://old.uk.fme.vutbr.cz/kestazeni/MKP/prednaska1_mkp.pdf)>.
- [9] ANSYS [online]. Dostupné na internete: <<http://zsc.zcu.cz/sw/ansys.html>>.
- [10] Nux s.r.o.: ANSYS WorkBench [online]. Dostupné na internete: <<http://www.techsoft-eng.cz/produkty/ansys---technologie-workflow/platforma-ansys-workbench/>>.

# Hodnotenie ekonomickej efektívnosti bioplynovej elektrárne

<sup>1</sup>Alexander MÉSZÁROS, <sup>2</sup>Patrik JAŠŇÁK, <sup>3</sup>Ján ZBOJOVSKÝ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>alexander.meszaros@tuke.sk, <sup>2</sup>patrik.jasnak@student.tuke.sk,  
<sup>3</sup>jan.zbojovsky@tuke.sk

**Abstrakt** — Bioplynové stanice predstavujú vhodný decentralizovaný zdroj z hľadiska stability dodávok elektrickej energie a spätných vplyvov na elektrizačnú sústavu. Kombinovaná výroba elektriny a tepla môže konkurovať konvenčným zdrojom z hľadiska efektivity vynaložených investičných nákladov a zaručiť tak trvalo udržateľnú nízkouhlíkovú elektroenergetiku.

**Kľúčové slová** — biomasa, bioplynová stanica, obnoviteľné zdroje energie, ekonomická efektívnosť

## I. ÚVOD

Problematika obnoviteľných zdrojov energie je aktuálnou a často diskutovanou témou v odborných kruhoch. Existuje mnoho aspektov prečo sa usilovať o rozvoj elektroenergetiky smerom k obnoviteľným zdrojom.

Jedným z hlavných dôvodov prečo sa spoločnosť uberať smerom k obnoviteľným zdrojom energie je fakt, že zásoby fosílnych palív sú obmedzené. Faktom je tiež, že bez výroby elektrickej energie v dostatočnom množstve by každodenný život nemohol fungovať tak ako ho dnes poznáme, pretože sme silne závislí na kontinuálnych dodávkach energií. Ak jedného dňa dôjde k vyčerpaniu zásob fosílnych palív ľudstvo musí byť pripravené na plynulý prechod k trvalo udržateľným obnoviteľným zdrojom energie. Pre znížovanie závislosti na fosílnych zdrojoch v budúcnosti je nutné hľadať riešenia už v súčasnej dobe. Každý svojím nápadom, víziou či postojom môže prispieť k rozvoju obnoviteľných zdrojov a ich uplatneniu v praxi.

Na druhej strane existuje mnoho prekážok, ktoré bránia rozvoju využitia týchto zdrojov, či už legislatívnych, ekonomických alebo technických. Negatívom môže byť napríklad ekonomická náročnosť výstavby projektov obnoviteľných zdrojov energie (OZE), ktorá často krát prevyšuje výhody plynúce z ich realizácie. Preto je nutné hľadať také technicko-ekonomické riešenie, ktoré minimalizuje negatíva.

## II. TECHNICKY VYUŽITELNÝ POTENCIÁL OBNOVITEĽNÝCH ZDROJOV ENERGIE V SR

Technicky využiteľný potenciál obnoviteľných zdrojov možno chápať ako využiteľnú kapacitu z celkového potenciálu, ktorú po zavedení dostupných technológií možno energeticky zhodnotiť. Limitujúcim faktorom môžu byť legislatívne, administratívne a environmentálne bariéry. Najväčší potenciál spomedzi všetkých OZE na Slovensku má energia z biomasy. Biomasa predstavuje až 44% technicky využiteľného potenciálu [5]. Územie SR má tiež vhodné podmienky na využívanie hydroenergetického potenciálu vodných tokov. Tabuľka 1 uvádza technicky využiteľný potenciál jednotlivých OZE v SR [3].

Zdroje s nestabilnou výrobou elektrickej energie sa považujú za nepredikovateľné, znamená to, že sú závislé na poveternostných podmienkach. Medzi tieto zdroje patria veterné a solárne zdroje. Svojím charakterom výroby zaťažujú elektrizačnú sústavu v dôsledku neočakávaných tokov elektrickej energie [4]. Ďalšou nevýhodou sú pomerné vysoké investičné náklady a dlhšia doba návratnosti investície a v kombinácii závislosti na slnečnom svite či vetre to môže byť prekážkou pri rozvoji malých fotovoltaických či veterných elektrární z hľadiska investora, pretože nestabilita budúcich príjmov znamená určité riziko pre investora. Geotermálna, vodná energia a energia z biomasy predstavujú stabilné zdroje, ktorých využitie na výrobu elektrickej energie by nezaťažovalo ES SR[5].



Tabuľka 1  
Technicky využiteľný potenciál OZE v SR [3]

Druh	Technicky využiteľný potenciál	
	GWh/rok	TJ/rok
Geotermálna energia	6 300	22 680
Veterná energia	605	2 178
Solárna energia	5 200	18 720
MVE	1 034	3 722
VVE > 10 MW <sub>e</sub>	5 573	20 063
Biopalivá	2 500	9 000
Biomasa	11 237	40 453
Spolu (bez VVE)	26 876	96 753
Spolu	32 449	116 816

Technicky využiteľný potenciál biomasy na Slovensku je 11 237 GWh/rok čo by pokrylo takmer 40% ročnej spotreby elektriny na Slovensku. Výhoda biomasy spočíva v nezanechávaní tzv. uhlíkovej stopy, pretože biomasa počas svojho rastu pri fotosyntéze spotrebuje oxid uhličitý, ktorý sa uvoľní do atmosféry pri jej energetickom využití. V podmienkach SR je možné energeticky zhodnotiť biomasu [3]:

- **lesnú biomasu** – palivové drevo, energetické porasty,
- **poľnohospodársku biomasu** – zvyšky z pestovania a spracovania plodín, záhradná biomasa zo sadov a viníc, výroba kvapalných biopalív (MERO, bioetanol), exkrementy hospodárskych zvierat, odpad z potravinárskeho priemyslu,
- **odpadovú biomasu** – odpad drevospracujúceho priemyslu, komunálny odpad, kaly z čističiek odpadových vôd.

Medzi pevné biopalivo patrí lesná biomasa, ktorú je možné priamo spaľovať alebo po spracovaní na drevnú štiepku spaľovať v automatizovaných kotloch. Ďalšou možnosťou je zhutňovanie drevej štiepky stlačením pri vysokom tlaku a teplote. Výsledným produktom sú drevené brikety alebo pelety, ktorých výhrevnosť je vyššia než výhrevnosť hnedého uhlia (14-16 MJ/kg) a pohybuje v hodnotách 17-19 MJ/kg.

Kvapalné biopalivá vyrobené z poľnohospodárskej biomasy sa primiešavajú do fosílnych palív (benzín, nafta) vo vhodnom pomere a znižujú tak emisie škodlivín v doprave.

### III. NÁVRH BIOPLYNOVEJ STANICE (BPS)

Aktuálny vývoj v poľnohospodárstve smeruje k diverzifikácii zdrojov príjmu v dôsledku nedostatočnej konkurencieschopnosti poľnohospodárskych družstiev. Primárne určenie poľnohospodárskej výroby sa transformuje z potravinárskej výroby na viacúčelovú teda aj nepotravinársku výrobu, čo zaručí udržateľný rozvoj poľnohospodárstva a vidieku. Práve rozšírenie činnosti poľnohospodárov o prevádzkovanie bioplynových staníc a o pestovanie energetických plodín ako zdroja pre tieto zariadenia je jednou z možností ako vylepšiť budúcu udržateľnosť poľnohospodárstva. Skúsenosti z Rakúska či Nemecka, kde intenzívne prebieha realizácia bioplynových staníc potvrdzujú pozitívny prínos pre vidiek a sú pre poľnohospodárov novým a stabilným zdrojom príjmov, vytvárajú a stabilizujú pracovné miesta, produkujú ekologickú energiu a kvalitné hnojivo. Významne tak prispievajú k ochrane životného prostredia a navyše k energetickej nezávislosti krajiny.

Pre návrh bioplynovej stanice pre porovnanie sa uvažovali tieto varianty:

- Variant A – model BPS, ktorá produkované teplo využíva len na technologické účely, resp. vlastnú spotrebu – objekt v priestoroch poľnohospodárskeho družstva,
- Variant B – model BPS, ktorá zvyškové teplo využíva nielen na vlastnú spotrebu, ale odpadové teplo dodáva priemyselným prevádzkam na technologické využitie, či vykurovanie objektov – objekt v priestoroch priemyselného parku.

Nájsť vhodnú lokalitu pre výstavbu a prevádzku BPS nie je až také problematické ako je to v prípade iných obnoviteľných zdrojov, napríklad malej vodnej elektrárne či veternej elektrárne, kde sa z hľadiska umiestnenia existuje mnoho bariér. Vhodná lokalita môže byť kdekoľvek kde je možné získať povolenie na výstavbu od príslušných úradov, avšak pre maximalizáciu efektívnosti takéhoto projektu je vhodné využiť už existujúcu infraštruktúru. Vhodne navrhnutá poloha môže maximalizovať všetky výhody plynuce z realizácie projektu BPS.

Z hľadiska efektivity využitia infraštruktúry je potrebné zaistiť kontinuálny prísun vstupov

(kukuričná siláž, hnojovica hospodárskych zvierat, močovka či iný biologický rozložiteľný odpad) a tiež vyvedenie elektrického výkonu a tepla. V prípade zabezpečenia kontinuity vstupných surovín je to možné vyriešiť vhodným dimenzovaním skladovacích priestorov. V prípade vyvedenia elektrického výkonu do sústavy to nie je až také jednoduché, pretože si investor nemôže vybrať miesto pripojenia do DS. Vhodné miest pripojenia z hľadiska konfigurácie siete určí PDS až po podaní žiadosti o pripojenie zdroja na výrobu elektriny do DS. Efektívnosť využitia zvyškovej tepelnej energie tiež závisí od vzdialenosti miesta spotreby.

Navrhovaná BPS podľa variantu A bude prevádzkovaná na báze mokrej fermentácie spracovaním kukuričnej siláže a hnojovice hospodárskych zvierat. Prevádzka BPS bude spracovávať tiež vlastnú poľnohospodársku produkciu. Výsledným produktom je výroba digestátu – vysokohodnotného hnojiva, výroba elektrickej energie a jej dodávka do elektrickej siete a výroba tepla pre vlastnú spotrebu. Vstupnou surovinou do BPS môže byť tiež odpad živočíšneho a rastlinného pôvodu a bioodpad z odrezkov drevín a kosenia trávy resp. iné využiteľné odpady z priemyselnej výroby. Uvažuje sa s obsluhou jedného pracovníka, ktorého náplňou práce bude 2x denne vložiť substrát do dávkovača a vykonávať dozor počas pracovnej doby. Prevádzka bude automatická a monitorovaná centrálnym riadiacim systémom.

S navrhovaným umiestnením druhého variantu sa uvažuje v priemyselnej zóne, kde by sa využila už vybudovaná infraštruktúra. Na hranici zóny sa nachádzajú vzdušné vedenia 110kV a 22kV, 6 trafostaníc a elektrická stanica 110kV/22kV. Predpokladá sa s využitím zvyškového tepla pre okolité prevádzky, ktoré ho môžu využiť pri svojich technologických procesoch, pre vykurovanie objektov alebo výrobu chladu formou trigenerácie. Priemyselná zóna sa nachádza v oblasti zameranej na poľnohospodársku výrobu čo znamená, že by bolo zaručené dostatočné zásobovanie vstupnými surovinami.

Požadované množstvo potrebného bioplynu pre chod zariadenia sa určí podľa normovanej spotreby bioplynu pre kogeneračnú jednotku (KGJ) o výkone 1 MW a je uvedené v tabuľke 2, základné technologické parametre BPS v tabuľke 3.

Tabuľka 2  
 Porovnanie množstva vstupných surovín pre variant A a B [2]

Variant	Vstupná surovina	Množstvo t/r	Výťažnosť m <sup>3</sup> /t	Produkcia m <sup>3</sup> /r
A	Kukuričná siláž	15 500	200	3 100 000
	Kravský hnoj	4000	80	320 000
	Prasacia hnojovica	2000	36	72 000
	Celkové množstvo produkcie bioplynu			3 492 000
B	Kukuričná siláž	17 000	200	3 400 000
	Odrezy cukrovej repy	1675	90	150 750
	Celkové množstvo produkcie bioplynu			3 550 750

Tabuľka 3  
 Základné technologické parametre BPS [2]

Základné parametre BPS		Variant A	Variant B
Inštalovaný výkon	elektrický	kW	1 000
	tepelný	kW	1 340
Produkcia bioplynu	m <sup>3</sup> /deň	10 080	10 080
Využitie zariadenia	h/r	8 250	8 250
	%	94,17	94,17
Vlastná spotreba BPS	%	4	4
Dodávka elektrickej energie	MWh	7 933	7 933
Výkupná cena elektriny	€/MWh	107,53	107,53
Technologická spotreba/dodávka tepla	%	15	15
Dodávka tepla	MWh	-	9 872
Cena tepla	€/MWh	-	49
Celkové vyrobené teplo	MWh	11 055	11 352

Vyprodukovaná tepelná energia bude slúžiť jednak na vlastnú spotrebu procesu fermentácie (variant A), a na predaj (variant B). Podľa zákona č. 309/2009 o OZE a KVET je kladená podmienka získania plnej výšky výkupnej ceny z elektrickej energie. Výrobca elektrickej energie v BPS spaľovaním bioplynu vyrobeného anaeróbnou fermentáciou za túto elektrinu dostane 100% výkupnej ceny ak je technologická spotreba meraná – podiel využitej technologickkej spotreby tepla musí byť 25% z celkovej výroby tepla. Ak technologická spotreba nie je meraná - podiel využitej technologickkej spotreby tepla musí byť 15% z celkovej výroby tepla [6].

Bilancia vyrobeného množstva elektrickej a tepelnej energie obidvomi variantmi je v tabuľke

4. Variant B sa od variantu A líši v tom, že vyprodukované teplo sa využije nielen na vlastnú spotrebu, ale aj na predaj. Možnosť ekonomického zhodnotenia odpadového tepla je riešením ako maximalizovať ekonomický efekt zariadenia.

Tabuľka 4  
 Bilancia elektrickej a tepelnej energie variantov A, B [2]

	Variant A	Variant B
elektrický výkon KGJ	1 MW	1 MW
využitie maxima výkonu	8 250 h	8 250 h
množstvo elektriny na svorkách generátora	8 250 MWh	8 250 MWh
straty + vlastná spotreba elektriny (4%)	$8\,250 \times 0,04 = 330$ MWh	$8\,250 \times 0,04 = 330$ MWh
množstvo elektriny dodanej do siete	7 920 MWh	7 920 MWh
výkupná cena za elektrinu	107,53 €/MWh	107,53 €/MWh
ročné tržby za predaj elektriny	$7\,920 \times 107,53 = 851\,637,6$ €	$7\,920 \times 107,53 = 851\,637,6$ €
tepelný výkon KGJ	1,34 MW	1,376 MW <sub>t</sub>
množstvo vyrobeného tepla	11 055 MWh <sub>t</sub>	11 352 MWh <sub>t</sub>
spotreba technologického tepla (15%)	$11\,055 \times 0,15 = 1\,658,25$ MWh	$11\,352 \times 0,15 = 1\,703$ MWh
nevyužitie odpadové teplo	9 391,75 MWh	-
teplo dodané zákazníkovi	-	$11\,352 - 1\,703 = 9\,649$ MWh
výkupná cena tepla	-	49 €/MWh
ročné tržby za predaj tepla	-	$9\,649 \times 49 = 472\,811$ €
ročné tržby celkom	851 637,6 €	$851\,637,6 + 472\,811 = 1\,324\,448,6$ €

#### IV. HODNOTENIE EKONOMICKEJ EFEKTÍVNOSTI PROJEKTU

Na účely posúdenia ekonomickej efektívnosti projekty slúžia najčastejšie tieto ekonomické ukazovatele:

- čistý tok hotovosti po realizácii projektu na báze cash – flow (CF),
- diskontovaný tok hotovosti – diskontovaná hodnota cash – flow (DCF),
- doba návratnosti investície,
- čistá súčasná hodnota – NPV,
- vnútorné výnosové percento – IRR.

##### A. Investičné náklady projektu

Výška nákladov na realizáciu navrhovaných variantov bioplynových staníc sa bližšie špecifikuje podľa požiadaviek projektovej dokumentácie a úprav zadávateľa na základe údajov dodávateľa z cenovej špecifikácie. V prípade variantu B sú investičné náklady vyššie o 24 % kvôli nákladom na výmenníkovú stanicu tepla a teplovod. Predpokladaná výška investičných nákladov a ich rozloženie pre navrhované variantné riešenia sa nachádza v tabuľke 5.

Tabuľka 5  
 Rozloženie investičných nákladov variantov A, B v eurách [2]

	%	variant A	variant B
Celkové investície	100	3 200 000	3 980 000
Stavebná časť	45	1 440 000	1 791 000
Technologická časť	35	1 120 000	1 393 000
Inžiniering ( projektová dokumentácia, povolenia )	5	160 000	199 000
Zemné práce	10	480 000	597 000

Návrh financovania pre prvý variant uvažuje financovanie z vlastných zdrojov investora na začiatku výstavby o výške 30% a neskôr dofinancovanie z úverových prostriedkov 70% z investičných nákladov. Pre druhý variant je navrhované financovanie v pomere 40% vlastné zdroje a 60% cudzie zdroje, čiže úver z banky. Financovanie prostredníctvom bankového úveru má svoje opodstatnenie na základe stanovenia výšky výnosu z vlastného kapitálu. Čím vyšší je podiel cudzích zdrojov tým vyšší je výnos z vlastného kapitálu pri rovnakej úrokovej sadzbe.

Uvažuje sa s úrokovou sadzbou 5 % a časom splatenia úveru 10 rokov.

##### B. Prevádzkové náklady projektu

Pod prevádzkovými nákladmi rozumieme náklady, ktoré slúžia na chod zariadenia. Prevádzka zariadenia si vyžaduje trvalú pracovnú silu na manipuláciu s materiálmi. Navrhované zariadenia uvažujú s jedným zamestnancom, ktorého náplňou práce je manipulácia

so substrátom a odpadmi a tiež dozor nad prevádzkou zariadenia počas pracovnej doby. Ďalšími nákladmi sú servis a údržba KGJ a mechanických zariadení ako sú dopravníky, čerpadlá, miešadlá. Náklady na fond opráv slúžia ako rezerva v prípade neočakávaných výdavkov. Pre oba varianty sa uvažujú rovnaké ročné prevádzkové náklady v celkovej výške 60 000 eur na rok.

### C. Výpočet toku hotovosti

Pri určení cash flow v prípade oboch variantov sa uvažovalo so zrýchleným odpisovaním, kde jednotlivé položky boli rozdelené podľa odpisovej skupiny na technologickú časť (generátory, miešadlá, zdvíhacie a dopravné zariadenia) a stavebnú časť (budovy, sklady digestátu, fermentory, silážne žľaby) [2]. Technologická časť patrí do 3. odpisovej skupiny s odpisovaním po dobu 12 rokov a stavebná časť do 4. odpisovej skupiny s odpisovaním po dobu 20 rokov. Na základe daňových prázdnin sa prvých 5 rokov neuvažuje daň s príjmu. Do celkových nákladov boli zahrnuté prevádzkové náklady, náklady na palivo – substrát, odpisy a poplatok za prístup do siete (G-komponent). Pri určení celkovej ceny za vstupnú surovinu sa uvažovalo s nasledovnými cenami: cena kukuričnej siláže 25 €/t pre variant A a 35 €/t pre variant B, cena odrezkov cukrovej repy 25 €/t. Celková cena za vstupný substrát v prípade variantu A vo výške 465 tis. €/rok, v prípade variantu B vo výške 637 tis. €/rok.

Tabuľka 6  
 Ukážka výpočtu toku hotovosti pre variant A [2]

Rok		2014	2015	2016	...	2033
tržby	elektrina	0	851 638	851 638	...	851 638
	teplo	0	0	0	...	0
	celkom	0	851 638	851 638	...	851 638
náklady	OPEX	0	60 000	60 000	...	60 000
	substrát	0	465 000	465 000	...	465 000
	odpisy	0	294 748	263 247	...	7 200
	G-komponent	0	19 720	19 720	...	19 720
	celkom	0	839 468	807 967	...	551 920
zisk	základ dane	0	12 170	43 671	...	299 718
	daň z príjmu	0	0	0	...	65 938
	rozdiel	0	12 170	43 671	...	299 718
úmor úveru	0	224 000	224 000	...	0	
investície celkom	960 000	0	0	...	0	
CF bežný rok	-960 000	82 918	82 918	...	240 980	
kumulovaný CF	-960 000	-877 082	-794 165	...	1 914 455	
odúročiteľ	1,05	1	0,95238095	...	0,42	
diskontovaný CF	-1 008 000	82 918	78 969	...	100 132	
kumulovaný DCF	-1 008 000	-925 082	-846 113	...	671 693	

Tabuľka 7  
 Ukážka výpočtu toku hotovosti pre variant B [2]

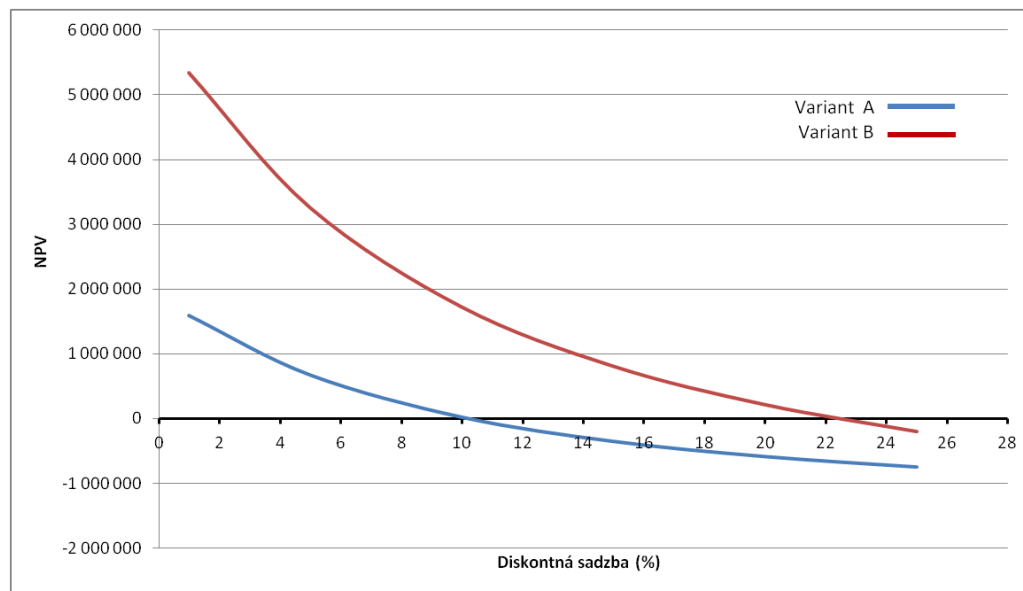
Rok		2014	2015	2016	...	2033
tržby	elektrina	0	851 638	851 638	...	851 638
	teplo	0	472 811	472 811	...	482 460
	celkom	0	1 324 448	1 324 448	...	1 334 098
náklady	OPEX	0	60 000	60 000	...	60 000
	substrát	0	636 875	636 875	...	543 500
	odpisy	0	205 633	382 964	...	17 910
	G-komponent	0	19 720	19 720	...	19 720
	celkom	0	922 228	1 099 559	...	641 130
zisk	základ dane	0	402 220	224 889	...	692 968
	daň z príjmu	0	0	0	...	152 453
	rozdiel	0	402 220	224 889	...	692 968
úmor úveru	0	238 000	238 000	...	0	
investície celkom	1 592 000	0	0	...	0	
CF bežný rok	-1 592 000	369 853	369 853	...	558 425	
kumulovaný CF	-1 592 000	-1 222 147	-852 293	...	8 370 792	
odúročiteľ	1,05	1	0,95	...	0,42	
diskontovaný CF	-1 671 600	369 853	352 241	...	232 037	
kumulovaný DCF	-1 671 600	-1 301 747	-949 505	...	4 989 776	

Na základe výšky čistej súčasnej hodnoty variantu A, ktorá je aj v prípade uvažovania diskontu väčšia než nula, možno projekt odporučiť na realizáciu. Na základe priebehu cash

flow možno určiť dobu návratnosti investície, ktorá je bez uvažovania činiteľa času na úrovni 13 rokov (hrubá návratnosť), v prípade uvažovania činiteľa času na úrovni 15 rokov (čistá návratnosť).

Podobne, aj v prípade druhého variantu je čistá súčasná hodnota väčšia než nula a odporúča sa realizácia projektu. Dobu návratnosti investície možno určiť na základe grafického zobrazenia cash flow je na úrovni 6 rokov a diskontovaná doba návratnosti na úrovni 7 rokov.

Závislosť na obr. 1 poukazuje na vplyv diskontnej sadzby na čistú súčasnú hodnotu, zároveň slúži na porovnanie navrhovaných variantov.



Obr.1 Závislosť NPV od diskontnej sadzby [2]

## V. ZÁVER

V návrhu variantných riešení sa uvažovalo s dvomi variantmi bioplynovej stanice. Hlavným rozdielom bola možnosť zhodnotenia odpadového tepla v prípade variantu B. Porovnávané varianty mali takmer identické technologické parametre, líšili sa len v investičných nákladoch, nákladoch na vstupné suroviny. Odlišnosti boli spôsobené rôznym umiestnením a charakterom prevádzky z pohľadu vstupných surovín, kde v prípade prvého variantu bolo plánované využitie odpadov z poľnohospodárskej výroby a v prípade druhého variantu bolo plánované využitie cielene pestovanej biomasy na energetické účely.

Bioplynové stanice predstavujú vhodný spôsob diverzifikácie energetických zdrojov a prispievajú tak k zníženiu závislosti na dovoze fosílnych palív. Prispievajú tiež k zlepšeniu regionálnej ekonomiky a zvyšovaniu pracovných miest. Avšak súčasne s rozvojom bioplynových staníc musí prebiehať aj rozvoj ES SR, pretože BPS predstavujú decentralizovaný zdroj, ktorý pôsobí určitými spätnými vplyvmi na sústavu, ktorá bola primárne určená pre centralizovanú výrobu elektrickej energie.

## POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva SR a Slovenskej akadémie vied projektom VEGA 1/0388/13.

## LITERATÚRA

- [1] Mészáros, A. "Ekonomika elektroenergetiky", Košice: TU-FEI, 2013. 117 s., ISBN 978-80-553-1512-6
- [2] Jašňák, P. "Analýza možností uplatnenia OZE v podmienkach ES SR", *Diplomová práca*. Košice: KEE FEI TU, 2014. 77 s.
- [3] Fábera, A. et al. "Atlas obnoviteľných zdrojov na Slovensku", Energetické centrum Bratislava, 2012. ISBN 978-80-969646-2-8
- [4] Program rozvoja SEPS, a.s. na roky 2013-2022. Dokument určený pre verejnosť. [cit 2014-3-20] Dostupné na internete: <[http://www.sepsas.sk/seps/Dokumenty/ProgRozvoj/2012/02/PR2022\\_verejnost\\_v2.pdf](http://www.sepsas.sk/seps/Dokumenty/ProgRozvoj/2012/02/PR2022_verejnost_v2.pdf)>
- [5] Koncepcia využívania OZE. [cit 2014-3-20]. Dostupné na internete: <<http://www.economy.gov.sk/koncepcia-vyuzivania-oze-5656/127341s>>
- [6] Vyhláška č. 372/2011 o spôsobe výpočtu ročnej výroby tepla pri výrobe elektriny spaľovaním bioplynu získaného anaeróbnou fermentáciou.

# Interactive Map of Campus for Android Operating System

<sup>1</sup>Peter LABUDA, <sup>2</sup>Ondrej KAINZ, <sup>3</sup>František JAKAB,

<sup>1,2,3</sup> Department of Computers and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>peter.labuda@student.tuke.sk, <sup>2</sup>ondrej.kainz@tuke.sk, <sup>3</sup>frantisek.jakab@tuke.sk

**Abstract** — The aim of this paper is to describe the creation process of the interactive campus map for Android devices. First, the analogical solutions are introduced. Next, the application itself is designed, taking into account given requirements. Developed application utilizes Google Maps. Application itself may be used as an aid for students within the campus of the university.

**Keywords** — Campus map, Interactive map, TUKE

## I. INTRODUCTION

The aim of this paper is to analyze the problems of developing applications for Android operating system and choose the appropriate tool for creating application, using GPS navigation system. The application will show the way in real time at the right place and will provide additional information about the campus.

The motivation for solving the problem of interactive campus map is that the campus is large and the people who are trying to find their way for the first time might have problem to find the desired place. Our solution is to create an application that will include a map without Internet connection. Internet connection will not be necessary and the search will be much faster. All the buildings on campus are stored in the database for effective handling. Internet connection will be required only for first run in order to download a map of the campus. Except for browsing the buildings on the map user will be able to view details of the building. An important part of the application is to provide navigation to a particular building. For its functionality, it is necessary to turn on the GPS system.

## II. RELATED WORK

Inspiration for developing campus application is drawn from analogous applications. The main screen will be similar to the start screen of application “University of Arkansas Campus Map [1]”, whose author is Kanat Bektemirov. On the start screen is a search box with the confirmation button and three tiles, Buildings, Departments and Dinning. Start screen can be seen in the picture (Fig. 1).

From Fig. 1, when clicked on button “Buildings” that is on the introductory page user is directed to the list that contains the abbreviation of the building and building full name. After clicking on an item in the list is displayed once again abbreviation of the building, phone number, full name of the building and two buttons: Show on MPA and Navigate. After clicking it is necessary to choose the application in which the action is to be carried out. Much more suitable would be if above mentioned functions were implemented in the application itself.

Viewing information on the map is well resolved in the application Tourist Attractions [2], which has a similar purpose; it shows tourist attractions of Slovakia, while enabling offline map usage. Attractions can be displayed in the list and sorted by distance from user. Third view is used to filter out the tourist attractions such as caves, castles etc., which simplifies the viewing on the map. To switch between the two views are used tabs. In the Fig. 1 and Fig. 2 are depicted screens of this applications.

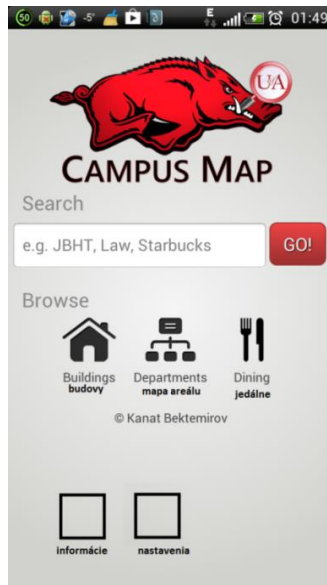


Fig. 1 Main activity of application Univers of Arkansas Campus Map [1].

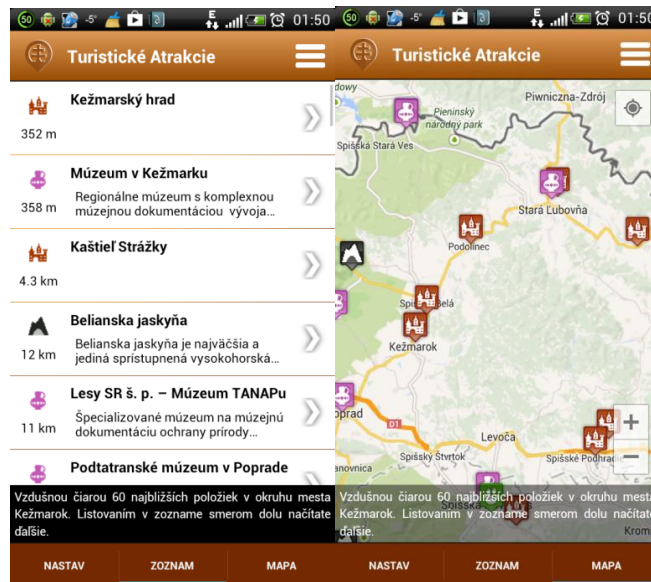


Fig. 2 Activities of application Turisticke atrakcie [2].

### III. PROPOSAL FOR APPLICATION SOLUTION

On the map will be shown buildings. Each building has its own unique markings and also the GPS position. The first step is to create a simple database of buildings for which we need just one table. Table will contain the following attributes:

- building\_id <pk> id of building,
- abbreviation abbreviation of the building, etc.: L9\_A, PK 6,
- type type of building, etc. library, dinner,
- description description of the building,
- lat latitude in decimal form numbers,
- lng longitude in decimal form numbers.

The next step is to display the map on which buildings will be displayed as stored in the database. Google Maps is utilized and adapts to the needs that are required by the application. For each building on campus is added a description of the building and the distance from it calculated using GPS [3].

Showing all buildings in the map at once is confusing when we are looking for one specific building or set of buildings with the same characteristics, e.g. dining room. Therefore we have proposed activity menu, wherein are two ways to search the building:

- Search box - will be used to search the building that contains some of their attributes (label, room, building type) searched part of the word.
- Filter Building - will be used to "filter out" searched buildings. For example, if a user wants to go to lunch and then go to coffee, mark tiles "dining room" and "café". Building stack updates after each click on the tile.

When user clicks on the "List" or "Map" searched or filtered out buildings will be displayed in a list or on a map. After choosing specific building detailed information in new activity are displayed, such as location, designation, name of the building and detailed information such as, opening hours and menu of food from the camera. This activity will also include the Navigate button which starts navigation to the selected building. The navigation activity will include user position showed on the map as a blue directional arrow, position of the target marked with green marker and colored line between these two points. Once the application starts the user has several options. The first is opening the activity menu. In this activity user searches user buildings from the database. These buildings are stored in the container of buildings. Stack is used for reading list activity and map activity. Stack of buildings is implemented Singleton design pattern that we ensure only a single instance of the stack and make it easily accessible to other classes (see Fig. 3).

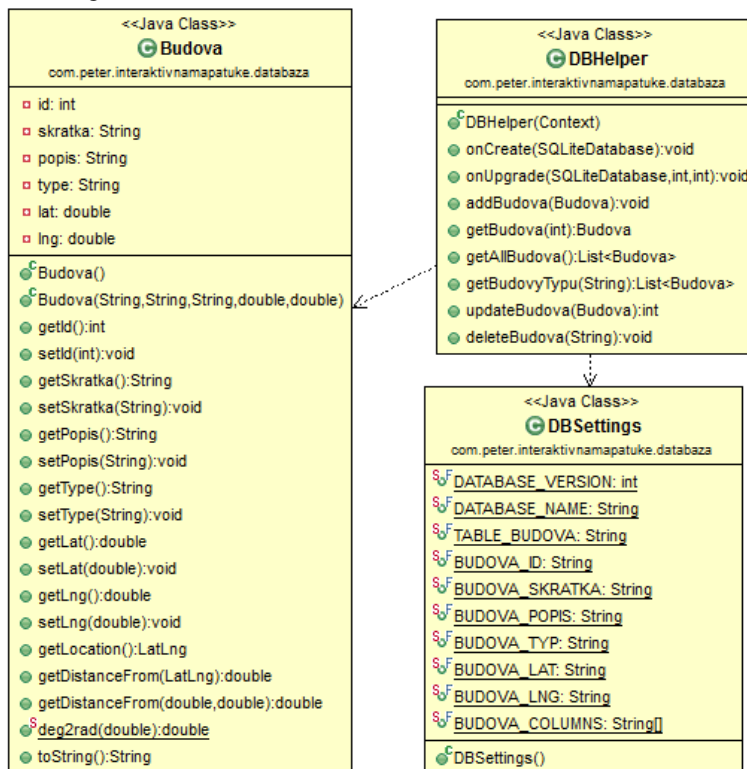


Fig. 3 Application class diagram.

A second use is to run the activity list. This activity shows all buildings of stack in the list view, this view contains information about building, such as location, building type, distance. Buildings from the stack are shown by distance or alphabetically in the list. Method for sorting can be set up in application settings.

A third option is to start activity MAP, wherein are displayed all the buildings of the stack as points on the map which will remain for the same information as if they were displayed in the list. User can zoom in or out, rotate map, etc. In Google maps can be easily displayed custom objects, calculated the distance, and used multiple modes of maps - for example classic map, a satellite map and a hybrid map. Even gestures may be used - scroll the map using one finger zooming maps shift two fingers together or from each other, rotate the map by rotation of two fingers on the screen etc. Therefore, it is the best to use Google Maps. In addition, it is not necessary to download new maps if the map is already saved map. From activities LIST or MAP (Fig. 4) user can switch to activity detail. In some dining rooms is a camera which scans the variety of meals. Showing menu of meals we have implemented in the activity detail. For showing the current menu it is necessary access to internet. Cameras are also available in the



University Library the COPYCENTRE, Jedlíková 5th and Jedlíková 9th. The camera can inform the user how many people are being served and the user can organize own time following this information. In activity detail some of the buildings are also linked to a web page. When clicking on such link user will be directed to Internet browser. From this activity can be accessed to the navigation activity as well.

Activity of Navigator shows our location and position of the building and between these points draws colored line. Launching this activity is allowed only if there is access to the Internet and GPS positioning system is running. At startup this activity appears on the map around the target location green marker and current location is marked with a red mark. When moving red mark will replace the mark in the form of a blue directional arrow on map shows the direction used is moving. The application offers choice of multiple languages. So far it is only translated into English, but the other languages are being implemented.

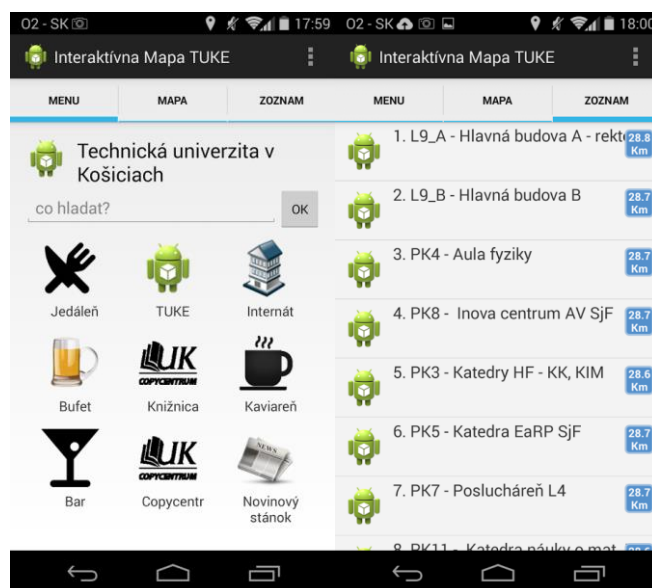


Fig. 4 Application layout.

#### IV. CONCLUSION

Developed application is easy and intuitive to handle. The main task of application is to display objects of campus and provide the most of useful information about them with the possibility to navigate to the selected school building. Application is not functional without at the least one-time connection to the Internet but we do not consider this to be major concern. Extension of in the future could be importing a calendar file from the academic system MAIS, which represents the timetable. The application would then know when the lectures begins and subsequently combine it with the navigational features.

#### ACKNOWLEDGMENT

Research described in the paper was supervised by Assoc. Prof. Ing. Frantisek Jakab, PhD., Computer Networks Laboratory at DCI FEEI TU in Košice. We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] Univers of Arkansas Campus Map [online] <https://play.google.com/store/apps/details?id=com.bektemirov.uark.campusmap&hl=en> (Accessed: 10 April 2014).
- [2] Turistické Atracie Slovensko [online] <https://play.google.com/store/apps/details?id=com.nano.android.turist> (Accessed: 22 April 2014).
- [3] Co to je GPS? Historie a úvod do problematiky [online] <http://www.svetmobilne.cz/co-to-je-gps-historie-a-uvod-do-problematiky/244> (Accessed: 10 January 2014).
- [4] R. Meier, *Professional Android Application Development*: Wiley Publishing, Inc., 2009. 409 s. ISBN 978-0-470-34471-2

# Interactive ISO-OSI Protocol Model and its Utilizing in Education

<sup>1</sup>Ronald TISCHLER, <sup>2</sup>Ondrej KAINZ, <sup>3</sup>František JAKAB,

<sup>1,2,3</sup> Department of Computers and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>ronald.tischler@student.tuke.sk, <sup>2</sup>ondrej.kainz@tuke.sk, <sup>3</sup>frantisek.jakab@tuke.sk

**Abstract** — The aim of this paper is to introduce an approach for design of a portal that is to be utilized as an educational tool. Features of proposed portal are Multilanguage support, modularity and user-friendly control. As a solution for implementation was chosen CMS Drupal. For the very same system custom visual module was developed and implemented. This module offers enhanced functionality for block positioning. Solution may be used in the real educational environment.

**Keywords** — CMS, ISO/OSI, MySQL, PHP, Protocols, TCP/IP

## I. INTRODUCTION

Since the dawn of computer there has been a necessity to solve the interconnection between networks. Initially, computer networks were used only to share data and peripherals, e.g. interconnecting of modems, printers etc. Today's computer networks offer bigger range of information services, without which we cannot imagine any activity on the computer. Nowadays internet is being used to collect and disseminate information, whether in the text form, pictures or even voice. To explain how networks work was created ISO/OSI model which describes communication between computers in seven layers, where each layer provides the functions necessary for the higher layer that utilizes the services of the lower layer. The issue of ISO/OSI model has already been solved many times all over the internet however there is no proper solution in Slovak language. Goal of this paper is to create a web portal which will provide such solution in both Slovak and English language, while enabling utilization of animations, which are to be created to help understand the issue. In addition, for people who wish to participate in the further development of the portal, changes are to be allowed with the approval of administrator. The CMS (Content Managements Systems) have expanded in recent years. Utilization of CMS system for this project is perfect solution. The CMS are stable, secure, and are a suitable solution for creating websites that fits our needs. Many web portals [1] are running CMS, for example portal of White House, Ubuntu or NASA academic program.

## II. ANALYSIS

### A. General description CMS (Content Management System)

CMS is a system for content management. Generally, it is an application which helps to create a website and manages its content structure. By content, we mean all types of documents, such as pages, images, structured data etc. The most important goal of any CMS is a strict separation of content, design and programming. There is no requirement for user to know how CMS is programmed. Similarly, one does not need the skills of web designer to know how to create new pages and arrange them in navigation of the website. Written content is automatically inserted into the interface created by the designer. Typically, CMS offers the ability to manage website content without knowledge of programming [2]. This system consists of files and database files – all of these provide features that can enable retrieval data from the database, whether it is content, pictures, files or anything else. The data obtained are packaged as HTML and are sent to the client. Idea is based on the fact that visitors are usually not interested how the site is made through they are more interested in the content and design of your site. CMS is running on the structure of a client – server (see Fig. 1). Server serves as your web space, where CMS is installed along with the database. Clients are represented by visitors of website itself [3].

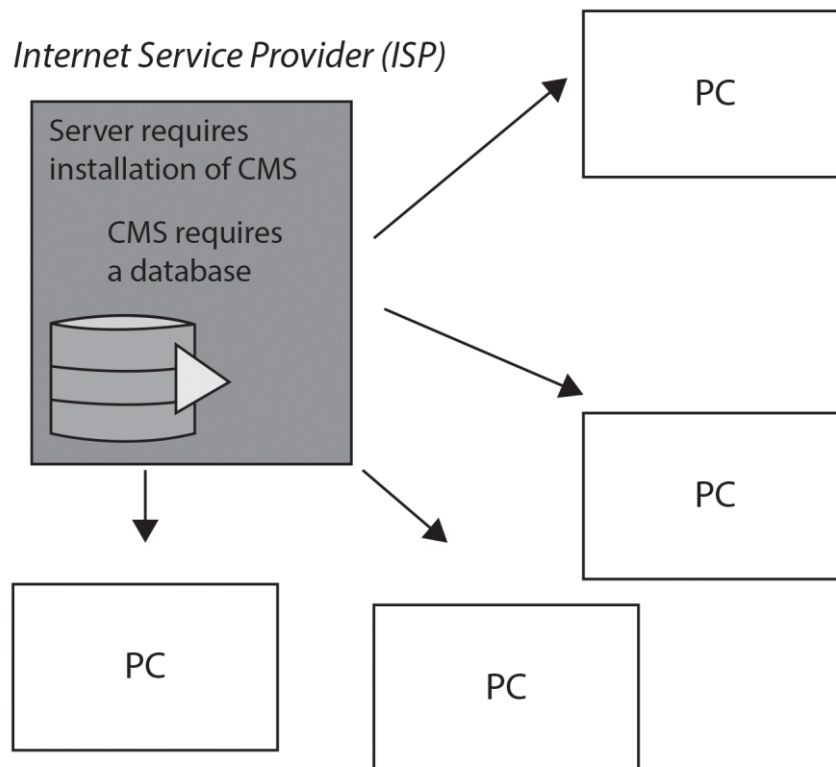


Fig. 1 CMS structure server-client.

### B. Security Solution of Users Rights

Initial step, before the proposal itself, is highly important – it represents dividing the rights of users who are to be registered on the portal. Rights of individual users can be for example [4]:

- Administrator – has the highest rights in the system and has absolute control. This role is able to make any changes to the system.
- Worker – has the second highest rights after administrator. This role can do everything without some special things for maintenance site.
- Registered user – can view and edit data after logging into the system. This role allows adding specific new topics to the forum or replying to previously existing.
- Anonymous user – have rights to view the contents of the portal, vote in pools and register.

### C. Program Applications

CMS does not require utilization of advanced tools for its management and modification. For example to edit themes is required only specific editor – in this project we used PSPad for PHP edit. In order to upload new modules to Drupal is required any simple FTP client, e.g. Total Commander. The last thing required is browser with HTML, CSS and JavaScript supports.

## III. SOLUTION AND RESULTS

The first step in the design of the portal is a need to design some basic interface of portal. Subsequently modules need to be implemented to enable enhanced features. Configuration and customization of portal to the correct form is very long process. In this chapter we presented implementation of created custom module.

### A. Solution of Portal

The portal is modular and easy to use, its visual appearance is implemented in an environment called Artisteer. In the Fig. 2 is depicted the home page – from the main menu at the top user is able to choose from network models and quiz, right above this menu is another dropdown menu that enables choice of language. As a default language on the portal was selected Slovak, since the presumption is that portal will be mostly utilized local users. On the very left side of the portal is the region in which is searching block, vertical navigation menu block, a module block showing the current online users, also block with statistics and block for registration or log-in. Adding of new topics is allowed only for registered users. In the middle of the page is a region which displays the content of the page and results of search. At the bottom of the page are less

important items. They are designed for a vote in polls, contact etc. Additional modules with expandable features such as online support, calendar, advanced forum and modules for increasing security against unwanted website registrations called spam were implemented on this portal as well. Security of portal is increased via regular updates of new version of CMS. These updates generally include new security patches and corrected security flaws.

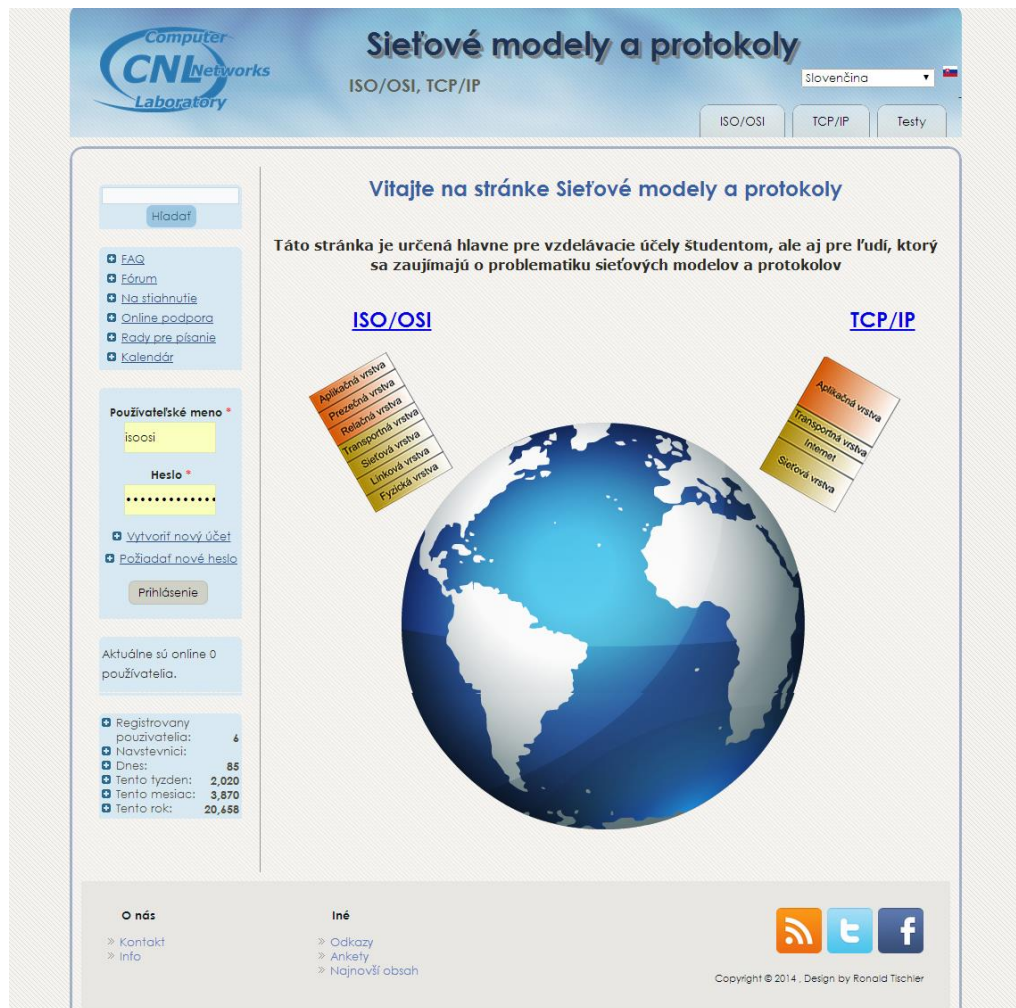


Fig. 2 Homepage of portal.

### B. Custom Visual Module

In the previous section was described a portal to which was applied developed custom module. The aim for developing this module is to facilitate the work with blocks for users of content management system Drupal 7. This module enables us to move each block up and down in the same region, among other regions, turn off block and move blocks between regions. Contextual links available in basic install of Drupal were utilized in order to achieve this functionality. Graphical representation of contextual links can be seen in Fig. 3.



Fig. 3 Contextual links in Drupal.

The module is created for Drupal version 7, and is not compatible with any other versions. The module depends on basic modules which are already included in each basic installation of Drupal 7. This module is not functional without basic modules block and contextual links. Graphical representation of the module is shown in Fig. 4.

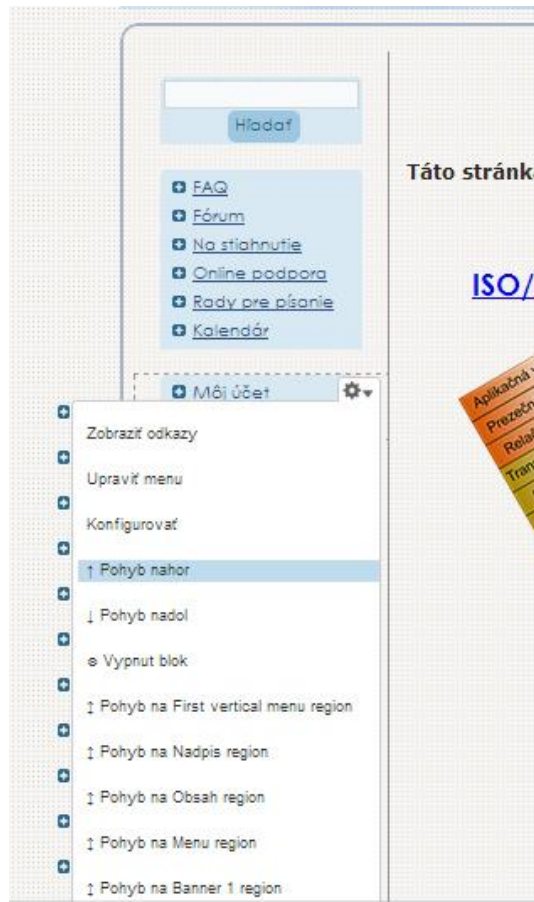


Fig. 4 Application of module.

### C. Evaluation of the Obtained Results

Portal is recommended to be used in the educational practice. Especially, it is suitable for students, educators however also for those people that are interested in network models and protocols. Designed portal is transparent, so it provides easy to use functionality.

Portal has the following features:

- Efficiency
- Modularity
- Extensibility
- Transparency
- Usability portal for educational purpose
- Multilingualism
- Security

Portal is created as a part of a complex project also contains elements as provided by from other members of team. Output from other members was interface for testing and animations for easier understanding the issue of network models and protocols. Modularity allows easy expansion of the portal which is to increase the qualitative features of the website. As the first of the possibilities for further development of the portal could be processing TCP/IP network model, since this model is not covered in detail, in this stage of project the portal is primarily focused on ISO/OSI network model. Another aim of the further development might be more detailed animations of specific protocols. Interesting solution could be authentication through LDAP protocol to interconnect all students from university with this portal.



#### IV. CONCLUSION

The main objective of this paper was development of a user friendly and modular portal focused on network models and protocols. Principal aim is to provide learning tool that will enable easy understanding and explanation of specific topic. The portal is suitable for people who are concerned with the issue of network models and protocols, however mostly it is aimed for students whose learning process may be enhanced utilizing this portal.

#### ACKNOWLEDGMENT

Research described in the paper was supervised by Assoc. Prof. Ing. Frantisek Jakab, PhD., Computer Networks Laboratory at DCI FEEI TU in Košice. We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. Also this work is the result of the project implementation: Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120020) supported by the Research & Development Operational Program funded by the ERDF (20%).

#### REFERENCES

- [1] Examples of web sites made with Drupal [online] <https://groups.drupal.org/files/drupal%20sites.html> (Accessed: 05 February 2014).
- [2] S. Hauschlidt, *CMS Made Simple 1.9 : Beginner's Guide*. Birmingham: Packt Publishing Ltd., 2011. 363 s. ISBN 978-1-849516-40-2.
- [3] L. Beighley, *Drupal for Dummies*. Hoboken: Wiley Publishing, Inc., 2010. 368 s. ISBN 978-0-470-55611-5.
- [4] The Drupal overview [online] <http://drupal.org/getting-started/before/overview> (Accessed: 20 March 2014).

# Meranie elektrofyzikálnych vlastností izolácie olej-papier

<sup>1</sup>Lukáš Lisoň, <sup>2</sup>Miroslav Kmec, <sup>3</sup>Iraida Kolcunová

<sup>1</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
Univerzita v Košiciach, Slovenská republika

<sup>2</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
Univerzita v Košiciach, Slovenská republika

<sup>3</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická  
Univerzita v Košiciach, Slovenská republika

[lukas.lison@tuke.sk](mailto:lukas.lison@tuke.sk), [miroslav.kmec@tuke.sk](mailto:miroslav.kmec@tuke.sk), [iraida.kolcunova@tuke.sk](mailto:iraida.kolcunova@tuke.sk)

**Abstrakt** — Transformátor hrá dôležitú úlohu pri poskytovaní spoľahlivej a efektívnej dodávky elektrickej energie a je jedným z najdôležitejších zariadení v prenose elektrickej energie a distribučných systémov. Papierová izolácia impregnovaná minerálnym olejom sa v transformátoroch používa už desiatky rokov. V dnešnej dobe je kladený veľký dôraz na používanie materiálov šetrných k životnému prostrediu. Prírodné estery sú považované za náhradu klasických minerálnych olejov. Tento článok je zameraný na výskum elektrofyzikálnych vlastností papierovej izolácie impregnovanej rastlinným repkovým olejom a minerálnom olejom. Na porovnanie vlastností papierovej izolácie impregnovanej prírodným esterom a minerálnym olejom boli vykonané merania stratového činiteľa a prierazného napätia.

**Kľúčové slová** — prírodný ester, minerálny olej, izolácia olej-papier, stratový činiteľ

## I. ÚVOD

Izolácia olej-papier je typ kombinovanej izolácie používanej v olejom plnených transformátoroch [1]. Väčšina vysokonapäťových transformátorov je plnená izolačným olejom, ktorý plní funkciu izolácie a tiež odvádza teplo [2]. V olejom plnených transformátoroch sa impregnovaný papier používa na izoláciu vodičov vinutí [5]. Najčastejšie používaným izolačným olejom v týchto zariadeniach vďaka svojej nízkej cene a vyhovujúcim vlastnostiam je minerálny olej [5]. Minerálny olej sa v týchto zariadeniach s úspechom používaný desiatky rokov s vynikajúcimi výsledkami [6]. Avšak využitie minerálneho oleja začína byť limitované v dôsledku zvýšenej pozornosti na environmentálnu bezpečnosť. V prvom rade minerálny olej nie je biologicky odbúrateľný a znečisťuje pôdu a vodu. V druhom rade minerálny olej sa vyrába z ropy, ktorá nepatrí medzi obnoviteľné zdroje a jej zásoby sa mňajú [2]. Prírodné estery sú izolačné kvapaliny šetrné k životnému prostrediu a sú považované za náhradu minerálnych olejov. Sú takmer úplne biologicky odbúrateľné a ich bod vzplanutia je vyšší ako 300°C. Dokážu viazať viac vody ako minerálne oleje [3]-[7]. Výskum vlastností prírodných esterov bude hrať v budúcnosti dôležitú úlohu, nakoľko už aj v súčasnosti sa v transformátoroch prírodné estery využívajú.

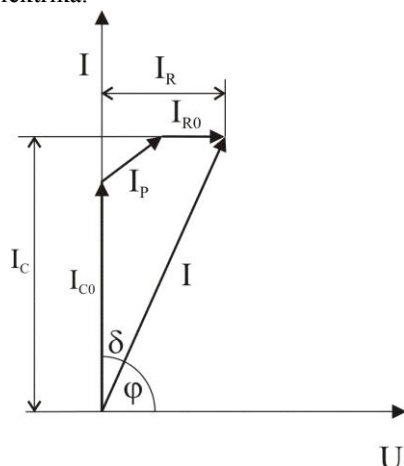
## II. ELEKTROFYZIKÁLNE VLASTNOSTI

Medzi základné dielektrické vlastnosti každého izolačného materiálu patrí stratový činiteľ, medzi vlastnosti patrí elektrické prierazné napätie a elektrická prierazná pevnosť.

### A. Stratový činiteľ

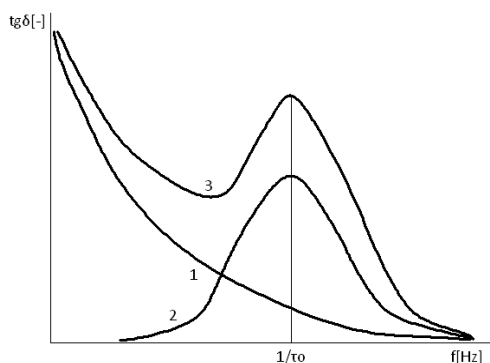
Stratový činiteľ vyjadruje mieru dielektrických strát a poukazuje na kvalitu izolačného materiálu. Čím je hodnota stratového činiteľa vyššia tým je kvalita daného materiálu horšia. V prípade merania stratového činiteľa, musí byť izolačný materiál zapojený ako dielektrikum kondenzátora. Izolácia olej-papier je zložená z rozdielnych druhov materiálu. Stratový činiteľ izolácie olej-papier závisí od stratového činiteľa každej zložky [3]. Na Obr.1 je zobrazený

fázorový diagram reálneho dielektrika.



Obr.1 Fázorový diagram reálneho dielektrika

$U$  predstavuje striedavé napätie priložené na dielektrikum. Prúd prechádzajúci dielektrikom je označený ako  $I$ . Kapacitná a reálna časť prúdu je reprezentovaná prúdmi  $I_C$  a  $I_R$ .  $I_p$  je polarizačný prúd a  $I_{R0}$  vodivostný prúd. Stratový činiteľ je závislý od frekvencie, napätia a teploty. Označuje sa  $\text{tg } \delta$  a je to bezrozmerná veličina, ktorá predstavuje podiel reálnej časti a kapacitnej časti prúdu. Závislosť stratového činiteľa od frekvencie je uvedená na Obr. 2. Krivka 1. predstavuje izolačný materiál s vysokou hodnotou vodivosti. Izolačný materiál s jedným polarizačným dejom je reprezentovaný krivkou 2. Frekvenčná závislosť reálneho dielektrika je kombináciou prvých dvoch prípadov a na Obr.2 je znázornená krivkou 3.



Obr. 2 Frekvenčná závislosť stratového činiteľa

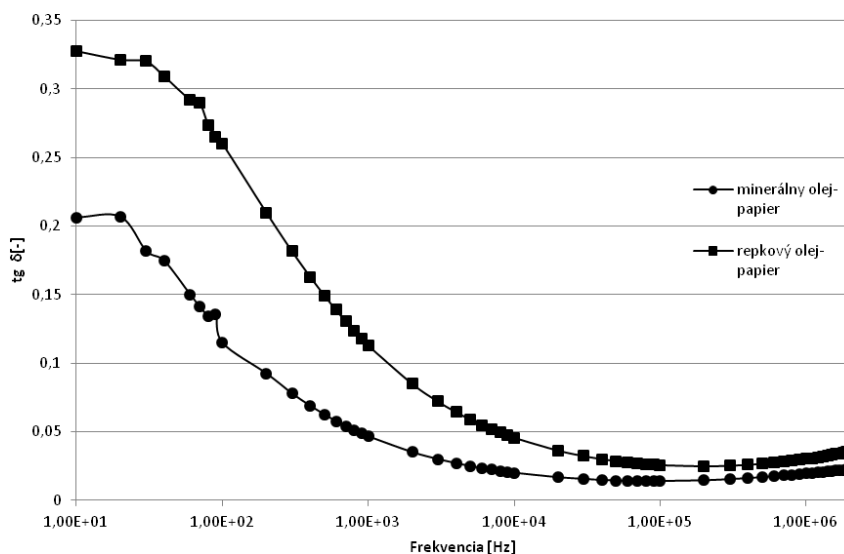
Na meranie frekvenčnej charakteristiky stratového činiteľa izolácie boli vybrané dva druhy kvapalných dielektrík a izolačný papier. Prvý typ izolácie je kombinácia repkového oleja a izolačného papiera, druhým je kombinácia minerálneho oleja Nynas-Lyra X a izolačného papiera. Pevná časť izolácie je tvorená šiestimi vrstvami papiera o hrúbke 0,06 mm. Celková hrúbka papierovej izolácie je 0,36 mm. Izolačný papier sa impregnoval rastlinným repkovým olejom a minerálnym olejom v nádobe po dobu 24 hodín. Po uplynutí uvedeného časového intervalu sú dutinky v papieri úplne vyplnené olejom, takže prítomnosť vzduchu je obmedzená. Stratový činiteľ uvedených vzoriek bol meraný pomocou RLC metra Agilent 4980a. Merací prístroj je zobrazený na Obr. 3, meria stratový činiteľ vo frekvenčnom pásme od 10Hz do 2 MHz.



Obr. 3 Digitálny RLC meter Agilent 4980a



Na Obr.4 sú uvedené namerané hodnoty stratového činiteľa papierovej izolácie impregnovanej prírodným repkovým olejom a papierovej izolácie impregnovanej minerálnym olejom Nynas-Lyra X



Obr. 4 Frekvenčná závislosť stratového činiteľa kombinovanej izolácie olej-papier

Výsledky ukazujú, že obidve vzorky majú podobný charakter frekvenčnej charakteristiky stratového činiteľa. Stratový činiteľ exponenciálne klesá s rastúcou frekvenciou. Izolačný papier impregnovaný repkovým olejom má vyššie hodnoty stratového činiteľa ako izolačný papier impregnovaný minerálnym olejom v celom frekvenčnom pásme. Podstatný rozdiel medzi hodnotami stratového činiteľa medzi jednotlivými vzorkami je vo frekvenčnom pásme od 10Hz do 0,1 MHz. Pri vyšších hodnotách frekvencie tento rozdiel nie je taký výrazný.

### B. Prierazné napätie

Prierazné napätie  $U_b$  je jedna z najdôležitejších vlastností izolačného materiálu [1].Predstavuje schopnosť izolačného materiálu odolávať elektrickému napätiu. Je to najnižšia hodnota elektrického napätia ktorá spôsobí vzrast elektrickej vodivosti na hodnotu vodivých materiálov. Prúd prechádzajúci miestom prierazu spôsobuje mechanické, tepelné a chemické zmeny dielektrika. Tieto zmeny sú tak výrazne že izolačný materiál nie je schopný obnoviť svoje pôvodné vlastnosti [9].Prierazné napätie papiera impregnovaného repkovým a minerálnym olejom bolo merané pomocou prístroja High-Voltage DTS-60D. Na Obr.5 je zobrazený merací prístroj na meranie prierazného napätia.



Obr. 5 High-Voltage DTS-60D

Na účely tohto merania boli použité tie isté vzorky izolačných materiálov ako pri meraní stratového činiteľa. Meranie prebiehalo pri striedavom napätí s frekvenciou 50 Hz. Napätie sa zvyšovalo z hodnoty 0 V s krokom 2 kV/s do prierazu. Izolačný papier bol umiestnený medzi elektródami pod hladinou oleja a hrúbka papierovej izolácie bola 0,36mm. Na každom type kombinovanej izolácie bolo vykonaných päť meraní. Vzorky papiera boli po každom prieraze

vymenené. Časový interval medzi jednotlivými meraniami bol 1 minúta. Výsledky merania prierazného napätia sú uvedené na Tabuľke 1.

Tabuľka 1 Prierazné napätie kombinovanej izolácie olej-papier

Vzorka	Up[kV]					Priemer
	1	2	3	4	5	
papier impregnovaný prírodným repkovým olejom	17,8	18	20,4	19,2	21,1	<b>19,3</b>
papier impregnovaný minerálnym olejom Nynas-Lyra X	18,6	18,5	18,9	17,1	18,4	<b>18,3</b>

Prierazné napätie prírodných esterov je vo všeobecnosti vyššie ako prierazné napätie minerálnych olejov. Prírodné estery dokážu viazať až tri krát viac vody ako minerálne oleje. Vlhkosť ktorá je obsiahnutá v papierovej izolácii je pohltená prírodnými esterami a tento fakt zapríčiňuje vyššie hodnoty prierazného napätia. Priemerná hodnota prierazného napätia repkového oleja a papiera je 19,3 kV a minerálneho oleja a papiera je 18,3 kV.

### C. Elektrická prierazná pevnosť

Elektrická prierazná pevnosť je popri stratovom činiteli a prieraznom napätí jedným zo základných kvalitatívnych charakteristík dielektrika. Izolačný materiál si uchováva svoje vlastnosti iba do určitej hodnoty intenzity elektrického poľa. Po dosiahnutí tejto hodnoty intenzity elektrického poľa rezistivita dielektrika rýchlo klesá na hodnotu rezistivity vodivých materiálov. V prípade homogénneho poľa je intenzita v medzielektródovom priestore rovnaká a preto elektrickú prieraznú pevnosť môžeme určiť zo vzťahu (1).

$$E_b = \frac{U_b}{d} \quad (1)$$

$U_b$  predstavuje prierazné napätie, ktoré sa udáva v kV. Medzielektródová vzdialenosť sa označuje ako  $d$  [mm]. Elektrická prierazná pevnosť papierovej izolácie impregnovanej rastlinným repkovým olejom a minerálnym olejom sa určí podľa vzťahu (1). V tabuľke 2. sú uvedené výsledky elektrickej prieraznej pevnosti.

Tabuľka 2 Elektrická prierazná pevnosť kombinovanej izolácie olej-papier

Vzorka	Ep[kV/mm]					Priemer
	1	2	3	4	5	
papier impregnovaný prírodným repkovým olejom	49,4	50	56,7	53,3	58,6	<b>53,61</b>
papier impregnovaný minerálnym olejom Nynas-Lyra X	51,7	51,4	52,5	47,5	51,1	<b>50,84</b>

Elektrická prierazná pevnosť dosahuje vyššie hodnoty pri papierovej izolácii impregnovanej repkovým olejom tak ako je to pri prieraznom napätí. Priemerná hodnota  $E_b$  pre izolačný systém tvorený repkovým olejom a izolačným papierom je 53,61kV/mm. Pre papier impregnovaný minerálnym olejom je táto hodnota 50,84kV/mm.

### III. ZÁVER

Tento článok prezentuje výsledky merania dielektrických vlastností papierovej izolácie impregnovanej prírodným repkovým a minerálnym olejom. Z nameraných výsledkov môžeme konštatovať nasledujúce závery:

- oba izolačné systémy majú podobnú závislosť stratového činiteľa na frekvencii. Stratový činiteľ exponenciálne klesá pri zvyšujúcej sa frekvencii. Papier impregnovaný repkovým olejom má vyššiu hodnotu stratového činiteľa ako papier impregnovaný minerálnym olejom v celom frekvenčnom pásme.
- prierazné napätie prírodných olejov je väčšie ako prierazné napätie minerálnych olejov. Priemerná hodnota prierazného napätia papierovej izolácie impregnovanej repkovým

olejom je 19,3 kV zatiaľ čo pri papierovej izolácii impregnovanej minerálnym olejom je táto hodnota na úrovni 18,3 kV.

- Tak ako pri prieraznom napätí tak aj pri elektrickej prieraznej pevnosti je táto hodnota vyššia u papierovej izolácie impregnovanej repkovým olejom. Priemerná hodnota  $E_b$  pre izolačný papier impregnovaný repkovým olejom je 53,64 kV/mm a pre izolačný papier impregnovaný minerálnym olejom je to 50,84 kV/mm.

Veľkou výhodou prírodných esterov je ich environmentálna nezávadnosť. Oproti minerálnym olejom majú však vyššiu hodnotu viskozity a to môžu spôsobiť ťažkosti v procese impregnácie.

#### POĎAKOVANIE

„Táto práca bola podporovaná Agentúrou Ministerstva školstva Slovenskej republiky pre štrukturálne fondy EÚ na základe projektu *Vývoj unikátneho nízkoenergetického statického zdroja pre elektrosystémy* (číslo projektu: 26220220029, prioritná os 2 Podpora výskumu a vývoja)“



#### IV. LITERATÚRA

- [1] J. Li, S. Grybowski, Y. Sun and X. Chen, “Dielectric properties of rapeseed oil paper insulation.” Annual Report Conference on Electrical Insulation and Dielectric Phenomena.2007.
- [2] B. Stacho, S. Rusek, R. Goňo, V. Král, T. Raška: Maintenance optimisation of transformers 110 kV/MV in distribution networks. In sborník konferencie EPE 2007, Ostrava:VŠB-TU Ostrava, 2007, 434-440, 7 s., ISBN 978-80-248-1391-2.
- [3] P. Kačor, Z. Hytka, D. Helštýn: CFD Analysis of Temperature Rise of Power Transformer, In Proceedings of the 14th International Scientific Conference Electric Power Engineering 2013, Kouty nad Desnou:VŠB-TU Ostrava, 2013, p.285-289, ISBN 978-80-248-2988-3.
- [4] J. Pihera, R. Polanský, P. Prosr, J. Ulrych: Dielectric Spectroscopy of Thermally Aged Insulation. In IEEE Conference on electrical insulation and dielectric phenomena. Piscataway: IEEE Service Center, 2012. s. 862-865. ISBN: 978-1-4673-1250-9 , ISSN: 0084-9162.
- [5] R. Lioa, J. Hao, G. Chen. Z. Ma, and L. Yang, “A Comparative Study of Physicochemical, Dielectric and Thermal Properties of Pressboard Insulation Impregnated with Natural Ester and Mineral Oil.“ IEEE Transactions on Dielectrics and Electrical Insulation Vol. 18, No. 5; October 2011.
- [6] J. Dai, Z. D. Wang,“ A Comparison of the Impregnation of Cellulose Insulation by Ester and Mineral oil“ IEEE Transactions on Dielectrics and Electrical Insulation Vol. 15, No. 2; April 2008.
- [7] J. Li, Z. Zhang, S. Grzybowski, and Y. Liu,“ Characteristics of Moisture Diffusion in Vegetable Oil-paper Insulation“ IEEE Transactions on Dielectrics and Electrical Insulation Vol. 19, No. 5; October 2012.
- [8] K. Marton, J. Balogh, J. Džmura, J. Petráš, “Elimination of dust particle sedimentation in industry environment“, Journal of Electrostatics, Volume 71, Issue 3, June 2013, Pages 208-213, ISSN 0304-3886, <http://dx.doi.org/10.1016/j.elstat.2012.12.026>.
- [9] M. Marci,“Meranie výbojovej činnosti v kombinovanej izolácii olej-papier,“diploma theseis, Košice,2009.

# Use of EMTP-ATP for Transients Calculation

<sup>1</sup>Marek KITZ, <sup>2</sup>Dušan MEDVEĎ

<sup>1</sup> Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>2</sup> Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>marek.kitz@student.tuke.sk, <sup>2</sup>dusan.medved@tuke.sk

**Abstract** — In presence, the electrical energy is one of the most important and most indispensable resources, but it can be influenced with many factors or phenomena, whether positively or negatively. One from such factors is transient. This paper demonstrates modeling and simulation of transients in ATPDraw program. All of the measured results, values, instruction texts to exercises, photos, downloadable files and calculations are stored in a database. The result of this paper is represented as a php web page, which loads the data from this database and runs the python scripts, which are responsible for calculation of electrical diagrams.

**Keywords** — Databases, EMTP-ATP, PHP, Python, Transient phenomena

## I. INTRODUCTION

Everybody of us has already met with electrical energy and its use. We know that it has some advantages and disadvantages. We talk mainly about its secondary phenomena, which does not give us the benefits. Transient phenomena have bad impact for our home appliances or machines and what is worse case that they have a bad influence on human's lives. Therefore it is very important to avoid them or we should restrict them to such an extent that they will hurt us or damage our appliances.

From the power engineering point of view, the main goal of this paper is to model and simulate the transient phenomena or steady-state with the help of ATPDraw program.

From the computer science point of view, these measured and simulated results we have to put in electronic version in the most appropriate way.

The main purposes of this paper are mostly education in power engineering area, explanation and description of steady-state and transient phenomena in practice, show the ways how to avoid of them, principle of creation the tool for demonstration of measured values in electronic form and creation of tool which is based on mathematical calculations of steady-state and transient phenomena.

After the combination of the mentioned purposes, we create very helpful tool for mentors, which they can use it for easier explanation of phenomena in power engineering.

## II. TRANSIENTS IN POWER ENGINEERING

We can say that transient is a phenomenon by which the electrical circuit changes the amount of accumulated energy non-periodically. Otherwise the energy of electric circuit is not changed or it changes periodically in steady state. The transient occurs in changing the structure or state of electric circuit. Under the term structure or state we understand connection or disconnection of electric supply, short-circuits, connection or disconnection of some part of electrical circuit and many others. Transients can happen only if electrical circuit contains accumulation element, inductor and capacitor are considered as accumulation elements. Only if these elements are used, the energy balancing can happened, at non-zero time. Otherwise transients cannot occur if there is a purely resistive circuit. [1] Transient phenomena are dividing into three groups:

### A. Transient wave phenomena

They are reflecting with the fastest changes in electric power systems. Their duration is very

short from microseconds to milliseconds. The process of changes is so fast, that we cannot neglect the velocity of spreading the electromagnetic waves in the individual parts of electric power systems. The example of this type is the atmospheric overvoltage.

#### B. Electromagnetic transient phenomena

They are slower than transient wave ones. Their duration is in tenths of seconds. We can neglect the spreading of electromagnetic waves, because of their smaller velocity. The example of this type are various types of short-circuits.

#### C. Electromechanical transient phenomena

The duration of these transients varies from tenths of seconds to tens of seconds. We can neglect the spreading of electromagnetic waves and also mostly electromagnetic inertia of parts. These types of transients are characterized by mechanical movement of rotor's generator. The example of this type is a power swing. [2]

### III. ATPDRAW PROGRAM

"ATPDraw is a graphical, mouse-driven preprocessor to the ATP version of the Electromagnetic Transients Program (EMTP) on the Microsoft Windows platform." [3] We used this program for modeling and simulation of steady-states and transients. We modeled examples which deal with DC circuits, one phase and three phase circuits, short-circuit currents, overvoltages, transposition, compensation and one exercise dealt with electronic in electric power engineering. We worked with many types of electrical components like resistors, inductors, capacitors, one- and three-phase's sources, one- and three-phase switches, ammeters, voltmeters, overhead lines, transformers, loads and many others. This program offers also the graphical representation of results via graphs. On the mentioned graphs we can demonstrate the changes which happened during the transients or steady-states. ATPDraw is also great tool for educational purposes, because via this program the beginners in power or electrical engineering can quickly learn the basics from many engineering areas. ATPDraw program has a Royal licence (for non-commercial use) and users can run it from Windows or Linux OS. [4] Our work was modeled in ATP Draw version 4.0.

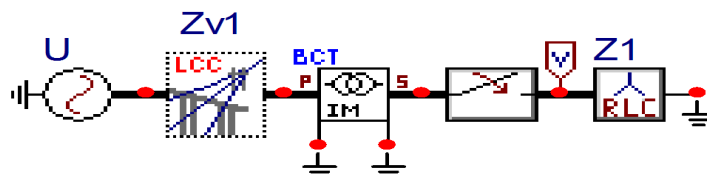


Fig. 1 Modeled circuit in ATPDraw program

### IV. CREATION PROCESS OF PRACTICAL EXAMPLES

After the modeling and simulation of examples, we created exercises from them, which are written as instruction texts. They are writing step by step and they fully describe the whole process of modeling and simulation. At the beginning of each practical example, there is an assignment of specific example and electric scheme. According this scheme, the electrical components are arranged. After the assignment, it follows its solution, which is detailed. Every electrical component is precisely described and we give the description of parameters, which are fulfilled and users will know what values are expected to be entered. Also mathematical solution is written in some types of electrical components, in order to know how we get the entered value. At the end of description of each electrical component, you can find the screenshot of its component windows, where you can see and control the entered values. If there are all electrical components placed and connected, we will get the final scheme. After that we have to fill important fields in menu *Settings*, then save our project and run its calculation. The user will choose the values which want to find out and press the button *plot* graph. The measured information is displayed in a form of graph. Theoretical conclusions and measured values in the tables are shown in chosen practical examples. The individual types of exercises are written in technical language.

### V. WEB DEVELOPMENT ENVIRONMENT

Whereas we want to show and present our results in the most appropriate way, we decided to demonstrate them as php web page, which is maintained through web development environment

called Wampserver. It is a freeware tool for creation and simple manipulation of our web applications. Wampserver consists of three important parts, which provide the proper running and management of our web applications.

#### A. Apache 2

Apache is the most popular web server. Its main goal is to provide and maintain HTTP server for various kinds of operating systems. The version 2.2.22 was used in our work.

#### B. MySQL

MySQL is open source database, which provides high-performance, scalable and web based database applications. It is also supported on common operating systems. The version 5.5.24 was used in our work.

#### C. PhpMyAdmin

It is responsible for simpler handling of MySQL over the web. It provides a GUI, where we can manage our databases in more simple way, but we can still have a possibility to directly execute the SQL commands. [5]

### VI. CREATION AND FULFILLMENT OF DATABASE

In phpMyAdmin menu, we created a database called “diplomovka”. After that we created its structure and we decided to create two tables. They are *web\_measurements* and *web\_navigation*.

#### A. *web\_measurements*

The table *web\_measurements* is the most important table in the database, because it contains the biggest amount of information. It consists of eight columns: *measurement\_id*, *measurement\_title\_eng*, *measurement\_title\_svk*, *measurement\_content\_eng*, *measurement\_content\_svk*, *navi\_id*, *measurement\_gallery\_svk* and *measurement\_gallery\_eng*. The *measurement\_id* column is responsible for numbering and identification of each record (measurement) which is stored in the database. Each measurement has its unique number. The second and third columns have the same goal. The columns, *measurement\_title\_eng* and *measurement\_title\_svk* contain the same title texts but in different languages. Also the fourth and fifth columns have the same goal. These columns, *measurement\_content\_eng* and *measurement\_content\_svk* contain the same texts in exercises but in different languages. These two columns contain the whole description of practical examples. They are the largest ones and they are written with the help of HTML tags. The column *navi\_id* groups the same types of practical examples and is used as separator in dropdown menu. The last two columns *measurement\_gallery\_eng* and *measurement\_gallery\_svk* contain the links for photos, which are stored in folders. The location and description of photos are written in links.

#### B. *web\_navigation*

The table web navigation consists of three columns: *navigation\_id*, *navigation\_svk* and *navigation\_eng*. The column *navigation\_id* is responsible for identification of each category in dropdown menu. The columns *navigation\_svk* and *navigation\_eng* contain the title names of categories in dropdown menu.

The data like images, installation files or documents are just written in a database via links, which contain the path to final destination. These data are stored in folders. Our work also contains such kind of folders, which are divided according their usage. These folders are stored in *X:\wamp\www\Diplomovka\_portal* directory.

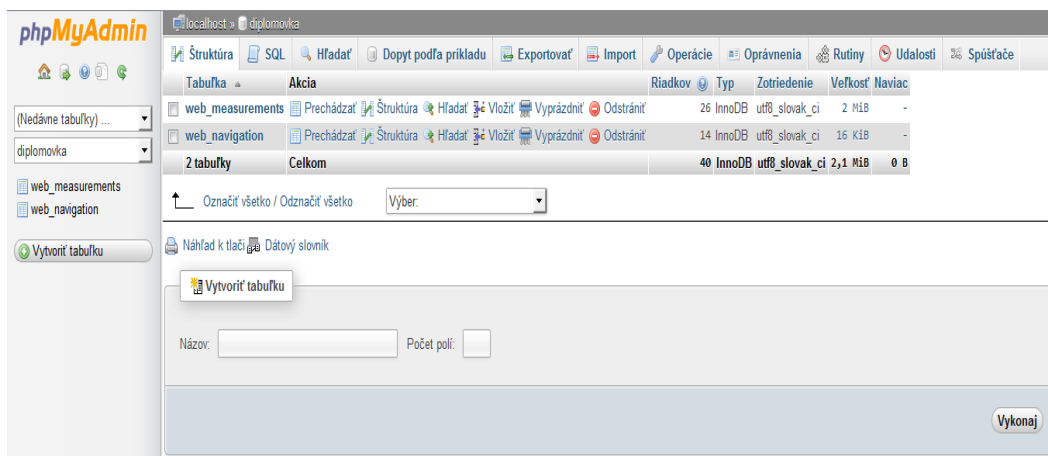


Fig. 2 Mentioned tables in “Diplomovka” database

## VII. PHP WEB PAGE

Free web template was used for creation of our php web page. Also for gallery, dropdown menu and graphs use free templates. We had to align all of mentioned templates with our components in this way that web page loads properly. The process of correction and modification of CSS styles is very difficult and time-consuming, because different types of web browsers behave different and the final look of web page differs among each other. We tried to find the best solution for some of web browsers.

If the web page is fully loaded, at the right upper side we can see the English and Slovak flag. They are responsible for changing the language on whole web page, if you click on them. Under them we can see the banner, where is written the name of department and logos of Technical University and department one. Under this banner we can see the drop down menu, from which we can go to different parts. It consists of fourteen main categories.

These are: *Home, DC Circuits, 1 Phase Circuits, 3 Phases Circuits, Short-circuit Currents, Overvoltages, Transposition, Compensation, Electronic in Power Engineering, Downloads, Gallery, Students, Theory and Calculations.* In these mentioned categories, we write the exercises' content, contact information about Ing. Dušan Medved', PhD., and important theory from power engineering. Also we can find there photos, installation files, guides, presentations or chosen examples which are calculated mathematically.



Fig. 3 Navigation menu on our web page

## VIII. PROGRAMMING

We divide the programming part into four main categories according the type of programming or scripting language on php codes, HTML codes, JavaScripts and python scripts.

### A. Php Source Codes

Php is an open source scripting language which is used for creation of web pages and also for general programming of various kinds of applications. [6] We decided to choose it, because nowadays php is very popular and many web pages are created in this programming language.

*Function.php* file contains two php functions. Function *db\_connection* is responsible for that php web page will load and take the data from database. Function *add\_j\_into\_imag\_part* checks the results, and according to mathematical sign of result, it will place “j” character in front of imaginary part. *Calculation.php* file contains the source codes where we assign values from URL parameters to variables, set the path to python binary and set the path where are scripts located. After that we called the scripts with *exec* function. Then it was necessary to parse the

results, because we needed to split real and imaginary part of results. At the end we put together real, imaginary part, HTML tags and units to write formatted results according the mathematical standards. Runge-Kutta example is not listed in *Calculation.php* file, because the results are not represented with numerical values, they are represented just in a form of graph. *Get\_graph.php* file contains all source codes which are dealing with setup and displaying of graphs on php web page. We created the legends and titles of axis for every graph. We also parsed results there, for real part and imaginary part. Then we specified which values will be used and displayed on  $x$  and  $y$  axis. After that we set the boundaries of graphs for  $x$  and  $y$  axis by getting the minimum respectively maximum of arrays. After that we set up the dimensions of graphs, scale and margin of graph. Then we set up the same parameters for  $x$  and  $y$  axis. We set up configuration of title axis, their font, margin, ticks and we can add the grid on graph. Then we created the lines which take values from arrays. At the end we created graph according the location and file name and this graph will be displayed on web page. *Calculation\_panels.php* file contains predefined values of individual examples and arrays which change the titles of examples according the language. In *Index.php* file there is the upper menu created, which contains the flags, logos of University and faculty and their names. Both of flags contain the hyperlinks and if you click on one of them, the whole page will be set on selected language. Then we have created the conditions according to buttons in dropdown menu were created and conditions of their behavior and appearance. At the first it was necessary to write *mysql\_query* function, which loads and selects our chosen data. We have created our conditions with the help of the columns *navi\_id* and *measurement\_id*. The principle of the conditions is in creation of categories respectively sub-categories in dropdown menu. After you click on main category and this category doesn't contain sub-category, load its content. But if category contains sub-category, show the list of sub-categories and after you click on sub-category the content will be loaded. Then we created conditions for gallery which appears on right side. The content of gallery will display just in a case of displaying practical examples. It will not display the full images, but thumbnails. So we had to set the path, position and name of the thumbnails. At the first we parsed the different images with delimiter “;”. The name and the path of image were parsed by delimiter “/”. At the end we had to divide the description of thumbnail according the language, = *\$lang* variable. So if the Slovak language is set, display “Obr.”, otherwise display “Fig.”. On our web page, there are applied some security checks. It means that if user changes the values of *category*, *example* or *language* in URL address on value which is incorrect or is not defined, it will load the home page.

### B. JavaScripts

Ajax is a collection of web development techniques, which allows you to change the content of web page without its reloading from the server side. [7] We decided to choose Ajax for the above mentioned reason in calculation of electric diagrams for displaying the mathematical results immediately, without reloading the web page. We have created the file called *script.js* where we placed all JavaScripts. The functions *calculateNew*, *calculateExample5* and *calculateExample6* are responsible for calculation of examples. They are composed from parameters which are indispensable for their running: *example* – number of exercise, *resultdiv* – container where result is written and *ID's* of input fields. If we got all necessary information we generated and added the button called *Generate Plot*, what basically mean that functions fulfill their roles and they can shift the information for plotting the graphs. If users don't write the correct values into the input fields, it displays the alert “All values must be entered correctly”. Under the term correct values we understand only numerical values and there can't be empty input field. The last two JavaScript functions *getGraph* and *getGraphRunge* are responsible for displaying the graph. Function *getGraph* contains three parameters: *example* – number of exercise, *exampleResult* – numerical result from example and *graphdiv* – container which contains the picture of graph. This function contains also part of code which deletes the previous graph if we want to simulate another state with different conditions. We created two seconds timeout for Ajax request what means that if the button *Generate Plot* is clicked, after the two seconds the graph will be displayed. Function *getGraphRunge* contains only two parameters: *example* and *graphdiv*. Their functionality is the same as in *getGraph* function.

### C. HTML codes

*Calculation\_panels.php* file also contains source codes for displaying the whole content of individual examples, if you click on example in dropdown menu *Calculation*. At the first, there are codes which are responsible for inserting of images, which show us how the example is



calculated numerically. Then there are created input fields with predefined values, which can be changed at any time. After that we created *Calculate* button which takes the known values, makes the calculation and writes the results. At the end we have created two containers or divs which insert their results to function *getGraph*. Runge-Kutta example doesn't contain the input fields with predefined values.

#### D. Python Scripts

We decided to choose Python scripts, because of simpler implementation of complex numbers, which are necessary in mathematical solution of electric circuits. Python offers us the every mathematical operation with complex numbers without that we had to derive it.

The first thing which we needed to make before the scripting was importing the modules: *math*, *cmath* and *sys*. After that we defined all of the variables. If all of variables were defined we started to create mathematical formulas. Since we had to deal with complex numbers, it was necessary to separate real and imaginary part into two variables. After that we had to create the result string to pass it into our php web page. There is an incomplete part of script which starts and ends with three dots. ... *return\_value=*str(real) + “,” *return\_value+=*str(P)+ “,” ...

The *Sympy* library was used for demonstration of Runge-Kutta method. So we had to install the *Sympy* module into the python folder. Then we imported library into our scripts as “*from sympy import \**”. The operations with matrices are demonstrated without complex scripting. Mathematical operations were provided simply just with the use of punctuation marks. The results of Runge-Kutta example are stored in *runge.txt* file. We wrote the script for creation of *runge.txt* file in these manner: *f = open('./runge.txt','w')* and then wrote the results in it: *f.write(return\_value)*. At the end we had to close the file: *f.close()*.

## IX. CONCLUSION

The goal of this work was to create educational tool which presents steady-states and transient phenomena in power engineering via ATPDraw program and the results are presented on php web page.

Our work has some advantages. The php web page is easy to navigate and the handling of web page is not so difficult. Users do not need to run some specific tools or programs to browse on the web page. Only thing which is necessary is web browser and Internet connection.

Like every other work, even this one has some disadvantages. Our version of ATPDraw program runs correctly only on 32-bits systems.

As an extension of this work we can suggest, the larger amount of exercises which will cover wider area of activity. Another suggested extension would be creation of forum, where users will publish their opinions and remarks. Academic communities would publish the articles or other publications, which will deal with the same or very relative issues or topics.

## ACKNOWLEDGMENT

This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences by the projects VEGA No. 1/0388/13.

## REFERENCES

- [1] BOJNA, Ivan: *Transient Response, Contact Resistance*. In: Electrical Engineering Terminology (EE) – technical paper for electrical and power engineering. Vol. 7, 2001, page 45-46.
- [2] MEDVEĎ, Dušan – HVIŽDOŠ, Marek: *Modelovanie v prostredí EMTP-ATP*. Košice: FEI TU Košice, 2011. 75s, ISBN 978-80-553-0776-3.
- [3] HOIDALEN, Hans Kristian: *New features in ATDraw ver. 4.0 for advanced simulation of electromagnetic transients in power systems*. October 2004, 58 p.
- [4] PR1KLER, László – HOIDALEN, Hans Kristian: *ATPDraw for Windows version 5.6 - User's Manual*. Trondheim: Norwegian University of Technology, November 2009, 270 p.
- [5] FOLTIN, Martin: *Prevádzkujeme Webserver na domácom počítači – WAMP*. In: posterus.sk. roč. 3, 2010, č. 6, ISSN 1338-0087.
- [6] PHP Manual: *Documentation – General Information*, [cit. 2014-2-12], Available on Internet:< <http://www.php.net/manual/en/faq.general.php#faq.general.what> >.
- [7] ULLMAN, Chris – DYKES Lucinda: *Beginning Ajax. Indianapolis*: Wiley Publishing Inc., 2007, 498 p. ISBN: 978-0-470-10675-4.

# Návrh a optimalizácia elektrických vykurovacích zariadení budov

Ján SABOL, Dušan MEDVEĎ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

jan.sabol.3@student.tuke.sk, dusan.medved@tuke.sk

**Abstrakt** — Predkladaný príspevok sa zaoberá návrhom elektrického vykurovacieho systému pre rodinný dom a športovú halu. Článok analyzuje správne rozloženie teplotného poľa a porovnanie investičných nákladov pre navrhnuté elektrické vykurovacie zariadenia.

**KLúčové slová** — elektrické vykurovacie systémy, podlahové vykurovanie, sálavé panely, konvektory

**Abstract** — The given article deals with the electrical heating system for the house and a sport hall. The article analyzes the temperature distribution and comparison of investment costs for the proposed electrical heating devices.

**Keywords** — electric heating systems, electric floor heating, radiant panels, electric convectors

## I. ÚVOD

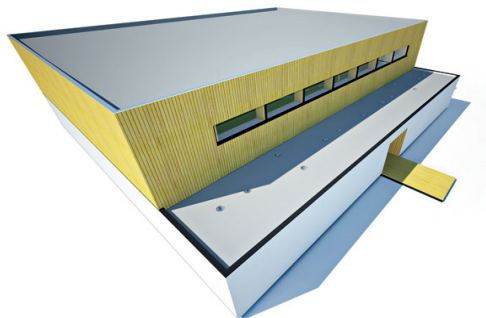
Súčasný trend neustáleho zvyšovania požiadaviek na vykurovanie obytných budov núti projektantov k presnejšiemu analyzovaniu teplotného poľa týchto objektov. Elektrické systémy majú výhody oproti neelektrickým v komfortnej regulácii, zníženej hlučnosti, či ľahkej údržbe.

Správny výber vykurovacieho systému je náročná problematika. Do úvahy je nutné brať niekoľko faktorov. Základnú úlohu zohrávajú vlastnosti budovy, ako sú napr.: konštrukcia, funkčné zameranie či geografické umiestnenie. V prípade, ak projektant pozná všetky vlastnosti budovy, môže vypočítať celkový tepelný príkon objektu (predstavujúci straty vo wattoch). Vypočítaná hodnota je príkon, ktorý musí dodať vykurovací systém.

Základom návrhu a optimalizácie je správne prerozdelenie príkonu využitím simulačného nástroja ANSYS. Vytvorenie simulačného modelu slúži na čo najefektívnejšie využitie vykurovacích systémov. Pomocou tohto nástroja je možné odsimulovať rozloženie teplotného poľa pre rôzne typy elektrických vykurovacích zariadení.

## II. VÝPOČET POTREBNÉHO TEPELNÉHO PRÍKONU

Vykurovacie systémy boli navrhnuté pre 2 objekty v 2 alternatívach: rodinný dom postavený v roku 1990 a v roku 2013 a samostatnú športovú halu pred a po rekonštrukcii. Objekty sa navzájom od seba odlišujú využitím stavebných materiálov a konštrukciou. Tepelný príkon bol vypočítaný podľa normy STN EN 12 831 [1].



Obr. 1 3D vizualizácia navrhnutého rodinného domu a športovej haly

### A. Rodinný dom

Materiál použitý na výstavbu budovy má významnú úlohu pri výpočte tepelných strát objektu. Novšia budova z roku 2013 mala obvodové múry postavené z materiálu s nižšou tepelnou vodivosťou, teda z materiálu s dobrou tepelno-izolačnou schopnosťou. Vypočítaný celkový tepelný príkon bol pre dom postavený v roku 2013 rovný 5936,05 W, pričom pre rodinný dom z roku 1990 je tepelný príkon 11 804,39 W.

### B. Športová hala

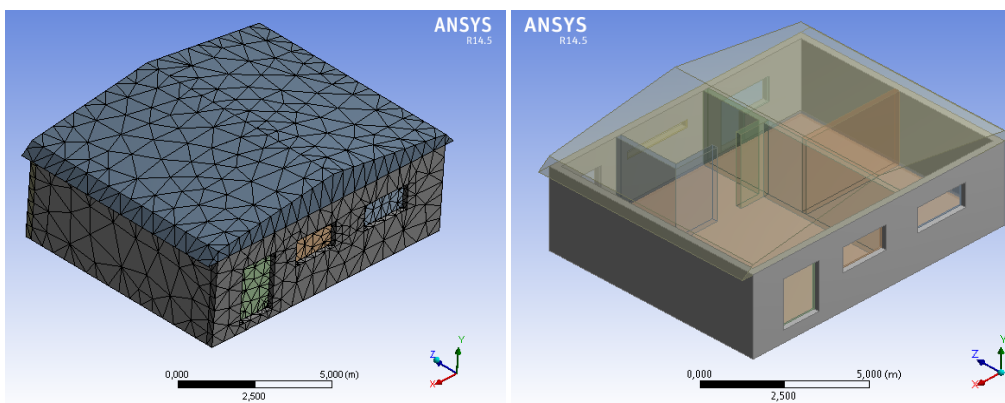
Pre rekonštruovanú športovú halu bola navrhnutá izolácia obvodových stien a výmena starých okien a dverí za nové, plastové. Touto zmenou (výmenou okien a dverí) sa znížil vypočítaný potrebný tepelný príkon z pôvodných 98 188,02 W na 55 892,43 W.

## III. VÝPOČET ROZLOŽENIA TEPLOTNÉHO POĽA

Pre zistenie rozloženia teplotného poľa je potrebné poznať hodnotu potrebného tepelného príkonu. Pomocou simulačného nástroja ANSYS je možné navrhnuť presný počet elektrických vykurovacích zariadení a ich správne umiestnenie na dosiahnutie potrebnej teplotnej pohody.

### A. Simulácia rozloženia teplotného poľa v okolí rodinného domu

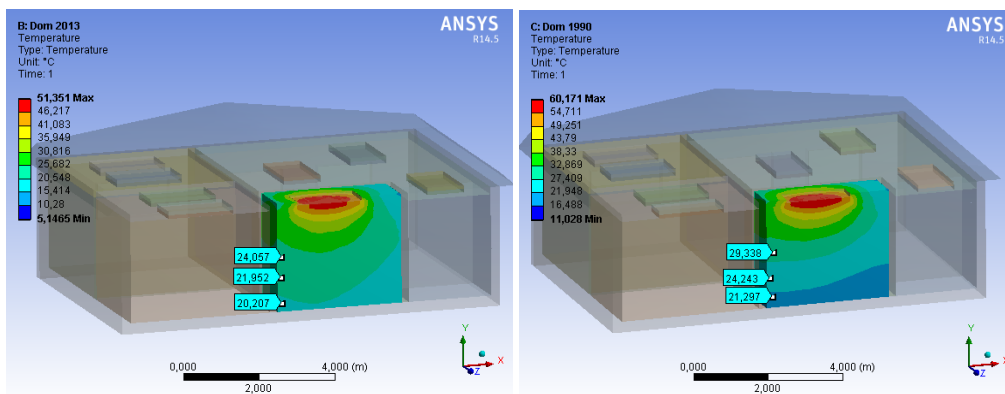
Zdrojmi tepla boli uvažované elektrické konvektory, sáľavé panely a podlahové vykurovanie. Na Obr. 2 je znázornený model domu s vysieťovaním a pohľad na jednotlivé miestnosti.



Obr. 2 Model rodinného domu a usporiadanie jednotlivých miestností

### Simulácia teplotného poľa pri vykurovaní sáľavými panelmi

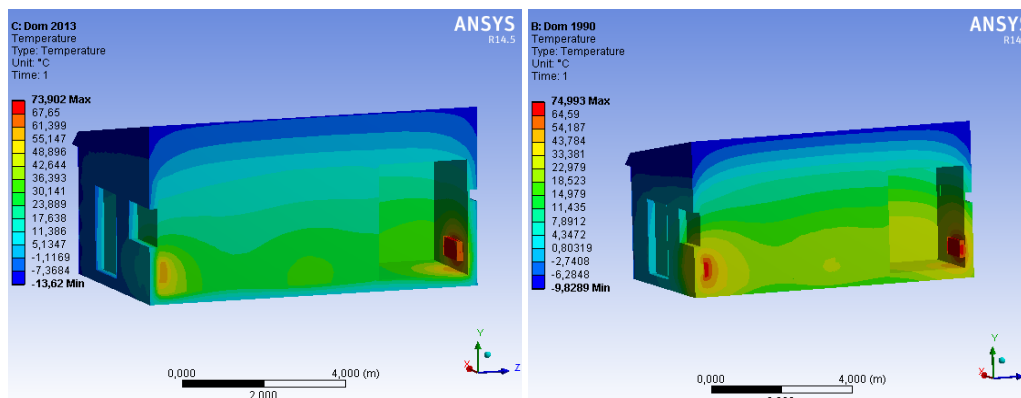
Na základe odporúčaní výrobcu boli sáľavé panely umiestnené na strope príslušných miestností (Obr. 3). Medzi strechou a sáľavým panelom je umiestnený tepelno-izolačný materiál. Na obrázku je znázornené rozloženie teplotného poľa v kúpeľni. Pri porovnávaní domu z roku 1990 a domu z 2013 je vidieť, že panel v kúpeľni 1990 musí mať vyššiu teplotu, aby dosiahol odporúčané teplotné podmienky pre kúpeľne. Pre dom z roku 2013 boli investičné náklady vyčíslené na 1665 € a pre dom z roku 1990 na 3414€.



Obr. 3 Rozloženie teploty v kúpeľni 2013 a v kúpeľni 1990

*Simulácia teplotného poľa pri vykurovaní konvektormi*

Pri šírení tepla konvekciou (prúdením) nastáva pod vplyvom gravitácie nerovnomerné rozmiestnenie teploty vo vnútri miestnosti. Na Obr. 4 je zobrazené rozmiestnenie teplotného poľa v hale a na stenách kúpeľne pre dom z roku 1990 a z roku 2013. Po porovnaní je možné usúdiť, že teplotné pole je rovnomernejšie rozložené v dome, pri ktorom boli použité stavebné materiály s nižším koeficientom tepelnej vodivosti. Vďaka nižším cenám konvektorov a ich jednoduchej inštalácii sú vstupné investície nižšie, ako pri sáľavých paneloch. Pre rodinný dom z roku 1990 boli investičné náklady vyčíslené na 1220 €, a pre dom z roku 2013 na 1040 €.

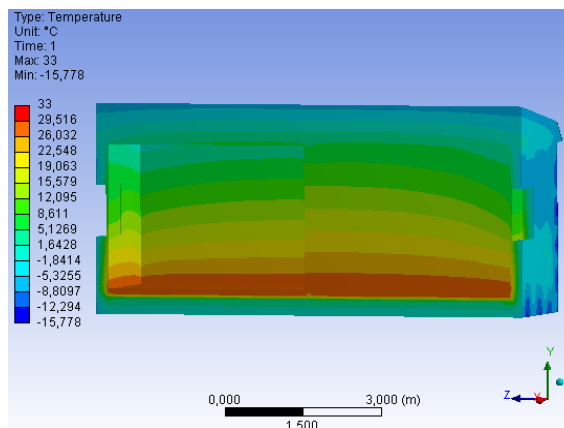


Obr. 4 Rozloženie teplotného poľa pre steny kúpeľne a vzduch v hale

*Simulácia teplotného poľa pri vykurovaní podlahovým vykurovaním*

Podlahové vykurovanie je možné realizovať pri inštalovaní termokáblov do podlahy. Pre obytné objekty je teplota podlahy 22 °C s výnimkou pre kúpeľne na 33° C. Pre zaistenie týchto teplôt miestností je problematické pre dom z roku 1990 (vysoké straty obvodovými múrmi). [3]

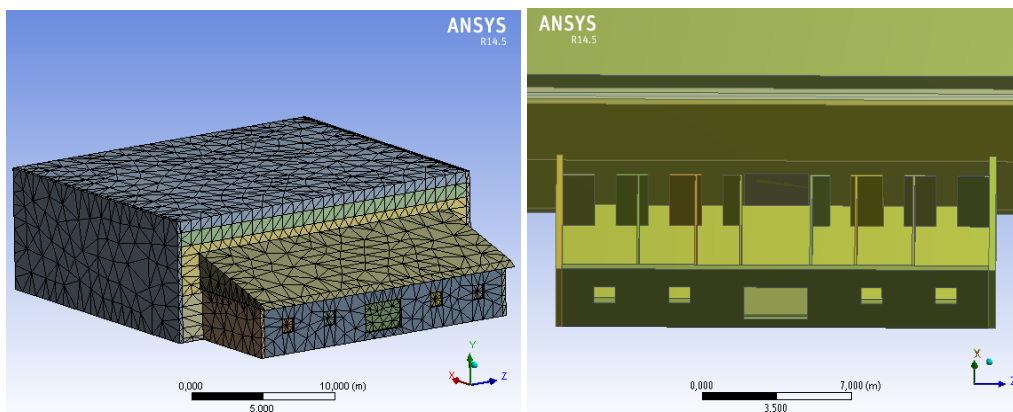
Riešenie rozloženia teplotného poľa pre dom z roku 2013 ukázalo, že podlahové vykurovanie má spomedzi vybraných elektrických vykurovacích zariadení, rovnomernejšie rozloženie teplotného poľa. Investičné náklady sú kvôli už vyššie zmieneným dôvodom dvojnásobné, oproti ostatným uvažovaným vykurovacím systémom. Vstupné investície pre elektrické podlahové vykurovanie rodinného domu postaveného v roku 2013 bolo vyčíslené na 3395 €.



Obr. 5 Rozloženie teplotného poľa pri podlahovom vykurovaní (spálňa a pracovňa)

*B. Simulácia rozloženia teplotného poľa športovej haly*

Návrh vykurovania športovej haly bol rozdelený na dva priestory – samostatne pre telocvičňu a šatne. Zo simulácií vyplynulo, že najvhodnejším elektrickým vykurovacím systémom pre telocvičňu sú sáľavé panely. Podlahové vykurovanie by nebolo schopné kompenzovať potrebný príkon a konvektory by neefektívne vykurovali celý priestor. Pre šatne sa uvažovalo s vykurovaním konvektormi a sáľavými panelmi.

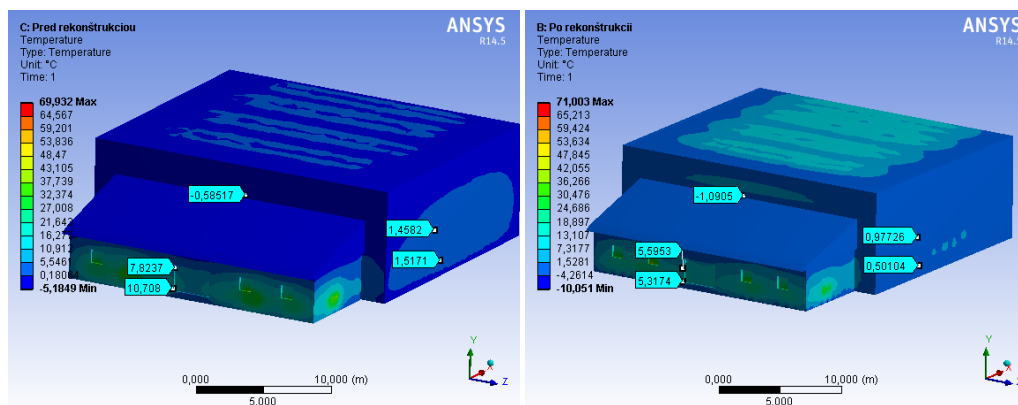


Obr. 6 Športová hala a rozdelenie miestností v šatniach

### Simulácia teplotného poľa pri vykurovaní sálavými panelmi

Sálavé panely sú vďaka svojmu umiestneniu na stropě schopné pri vyššom výkone zabezpečiť rovnomerné teplotné pole aj v hale pred rekonštrukciou. Ich investičné náklady (vrátane montáže) boli pre športovú halu vypočítané na 13 170 € pred rekonštrukciou a na 10 208 € po rekonštrukcii.

Rozloženie teplotného poľa športovej haly pred a po rekonštrukcii je uvedené na Obr. 7. Pre halu pred rekonštrukciou sú straty rovnomerne rozložené po obvode budovy. Po rekonštrukcii, respektíve zateplení, je obvodová konštrukcia budovy tepelne menej vodivá. Vyššie straty boli v okolí okien, dverí a strechy.

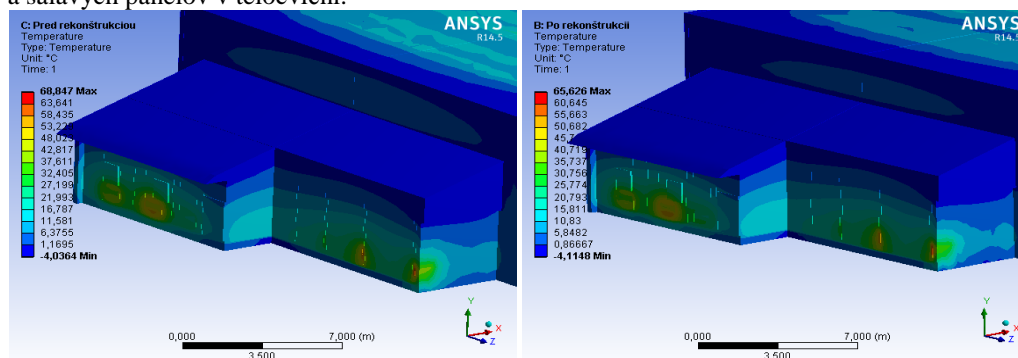


Obr. 7 Tepelné straty športovej haly pred rekonštrukciou a po rekonštrukcii

### Simulácia teplotného poľa pri vykurovaní konvektormi

Konvektory boli umiestnené na obvodové steny a pod okná. Takéto rozmiestnenie zabezpečilo rovnomerné rozloženie teplotného poľa v šatniach. Keďže sa jedná o športovú halu, je nutné si uvedomiť, že jej využívanie nie je neprerušované. Vo večerných hodinách, či pri zimných prázdninách, je hala len temperovaná. Pri vyčíslení vstupných investícií je potrebné brať do úvahy aj skutočnosť, že športová hala nie je využívaná neprerušovane.

Investičné náklady vykurovacieho zariadenia pre halu pred rekonštrukciou sú 8 550 €, po rekonštrukcii 6 900 €. Cena vykurovania zahŕňa montáž a kúpu konvektorov v šatniach a sálavých panelov v telocvični.



Obr. 8 Porovnanie rozloženia teplotného poľa v šatniach športovej haly pred a po rekonštrukcii

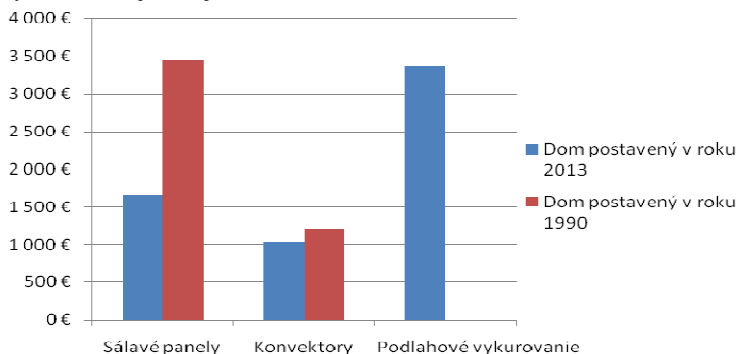
#### IV. POROVNANIE ELEKTRICKÝCH VYKUROVACÍCH SYSTÉMOV

Porovnanie uvažovaných elektrických vykurovacích systémov bolo rozdelené na dve časti. V prvej časti boli porovnávané hodnoty rovnomernosti rozloženia teplotného poľa získané simulovaním modelovaných objektov. Druhá etapa porovnávala vstupné investičné náklady.

##### A. Porovnanie elektrických vykurovacích zariadení pre rodinné domy

Pre rodinný dom, postavený v roku 1990, boli uvažované sálavé panely a konvektory, pričom z grafov rozloženia teplotného poľa sa sálavé panely javia ako efektívnejšie (rovnomernejšie rozloženie teplotného poľa). Uložením panelov na strop zostáva priestor na bočných stenách napríklad pre nábytok. Investičné náklady na kovektory sú nižšie.

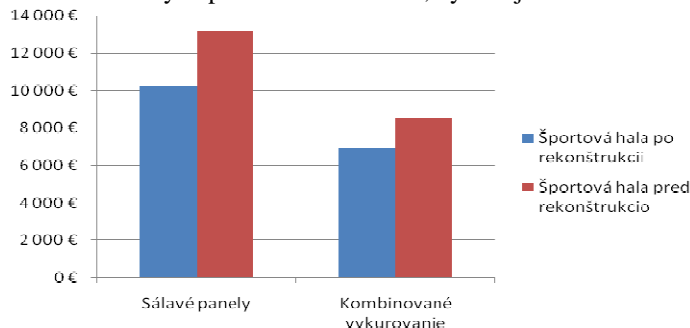
Dom postavený v roku 2013 má vďaka nižším tepelným stratám obvodovými múrmi možnosť využiť podlahové vykurovanie. Pri takomto vykurovaní je rozloženie teploty v miestnosti prirodzenejšie (tepelná pohoda). Investičné náklady sú však vyššie, ale životnosť podlahového vykurovania je dvojnásobná.



Obr. 9 Porovnanie investičných nákladov pre rodinné domy

##### B. Porovnanie elektrických vykurovacích zariadení pre športovú halu

Výber vykurovacieho zariadenia pre športovú halu je špecifickejší ako u rodinného domu. Kvôli nižším nárokom kladeným na vykurovanie počas zimnej sezóny je možné viac zohľadniť investičné náklady ako rozloženie teplotného poľa. Z porovnávaných systémov je kombinácia konvektorov v šatniach a sálavých panelov v telocvični, vyhovujúca.



Obr. 10 Porovnanie investičných nákladov pre športovú halu pred a po rekonštrukcii

#### V. ZÁVER

Tento príspevok mal za úlohu predstaviť porovnanie vybraných vykurovacích systémov pre rodinný dom a športovú halu (telocvičňu) v dvoch variantoch (pred a po rekonštrukcii). Využitím nástroja ANSYS je možné stanoviť rozloženie teplotného poľa daného objektu a následne sa rozhodnúť pre voľbu vhodného vykurovacieho systému.

#### POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied grantom VEGA č. 1/0388/13.

#### LITERATÚRA

- [1] MEDVEĎ, Dušan; NOVÁK, Pavel: *Premeny elektrickej energie*. ISBN 978- 80- 553-0883-8.
- [2] STN EN 12 831: 2003. *Vykurovacie systémy v budovách, Metoda výpočtu projektovaného tepelného príkonu*.
- [3] PETRÁŠ, Dušan, KOUDELKOV8, Daniela: *Teplovodní a elektrické podlahové vytápění*. Bratislava: Jaga group s.r.o., 2004. ISBN 80-88905-97-4.

# Analýza teplotného poľa elektroenergetických zariadení

*Maximilián BONDRA, Dušan MEDVEĎ*

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

maximilian.bondra@gmail.com, dusan.medved@tuke.sk

**Abstrakt** — Predložený príspevok sa zaoberá analýzou teplotného poľa vybraných elektroenergetických zariadení. V úvode sú opísané teoretické poznatky z termodynamiky, ako napríklad generovanie tepla, vplyv oteplenia a mechanizmy prenosu tepla. V ďalšej časti sú aplikované tieto poznatky na vybrané zariadenia využitím nástroja ANSYS.

**KLúčové slová** — transformátor, asynchrónny motor, dovolené oteplenie, dovolená teplota, teplo, ohraňovaná transformačná stanica

## I. ÚVOD

Prechodom elektrického prúdu vodičom sa vodič ohrieva a zároveň ohrieva aj svoje bezprostredné okolie. Každá teplota, ktorá prevyšuje pracovnú teplotu, či už hovoríme o vodičoch alebo elektrických strojoch, pôsobí degradujúco na izoláciu, vodiče a samotné zariadenie. V príspevku sa venujeme riešeniu rozloženia teplotného poľa v okolí asynchrónneho motora, transformátora a v okolí ohraňovanej transformačnej stanice.

Striedavý elektrický prúd vytvára v elektrických stojoch straty v elektrickom a magnetickom obvode. Vplyvom týchto strát sa generuje teplo, ktoré zvyšuje teplotu samotného zariadenia. Ak nie je zabezpečené dostatočné odvádzanie zvyškového tepla, môže to viesť k následnému poškodeniu, resp. zničeniu takýchto zariadení. Preto pri elektrických zariadeniach je stanovená hranica maximálneho oteplenia, ktoré sa nesmie prekročiť.

## II. MECHANIZMY PRENOSU TEPLA

Predpokladajme, že v určitom priestore je nehomogénne rozložená teplota, potom v danom priestore musí prebiehať proces prenosu tepla. Podľa druhej termodynamickej vety je tento proces uskutočňovaný v smere poklesu teploty, to znamená z teplejšej oblasti do chladnejšej oblasti. Tento proces výmeny tepla charakterizujeme vektorovou veličinou – vektorom hustoty tepelného toku  $q$ , ktorý predstavuje množstvo tepla v jouloch prechádza cez jednotku plochy [2].

Kvôli zjednodušeniu, je možné povedať, že prenos tepla je uskutočňovaný troma mechanizmami: *vedením* – kondukciou; *prúdením* – konvekciou; *žiarením* [2].

### A. Prenos tepla vedením – kondukcia

S prenosom tepla vedením sa stretáme pri ohrievaní tuhého telesa (napríklad kovovej tyče). Molekuly ohrievaného konca tyče sa okolo svojich rovnovážnych polôh pôsobením zdroja tepla rozkmitajú a postupne odovzdávajú energiu susedným molekulám až pokiaľ bude stredná kinetická energia molekúl v celom telese rovnaká. Keďže teplota sa definuje ako stredná kinetická energia molekúl, dochádza k postupnému ohrievaniu tyče od jedného jej konca po druhý. Za prenos tepla sa tu považuje taký prenos, pri ktorom molekuly telesa nemenia svoje polohy, nedochádza k pohybu telesa ako takého, nedochádza k makroskopickému pohybu [3].

### B. Prenos tepla prúdením – konvekcia

Pri zohrievaní kvapaliny v sklenenej nádobe si môžeme všimnúť stúpanie teplej kvapaliny nahor a klesanie chladnej kvapaliny nadol. Premiešavaním, ktoré je dôsledkom týchto pohybov, sa teplota v sklenenej nádobe vyrovnáva a aj keď sa teplo privedie zdola, zohreje sa aj



kvapalina pri hladine. Takémuto prenosu tepla, pri ktorom ide o makroskopický pohyb látky, t.j. prenos hmoty, hovoríme **konvekcia** [3].

### C. Prenos tepla žiarením

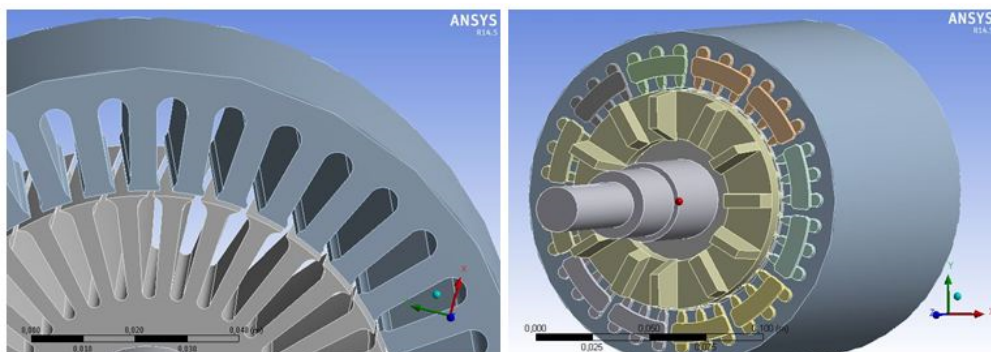
Slnko je od nás vzdialené milióny kilometrov, no aj tak môžeme cítiť, ako na nás žiari. Tepelná energia k nám prichádza zo Slnka cez riedke prostredie, teda v podstate bez látkového nositeľa. Tepelná energia slnečného žiarenia je energiou elektromagnetického poľa, na ktoré sa môžeme pozeráť ako na vlny, alebo ako na prúd fotónov. Žiarenie môžeme pozorovať takmer pri každom predmete, ktorý vyžaruje teplo, či sa jedná o vlákno žiarovky alebo o žeravé uhlie. Teplo žiari z každého telesa, ktoré má väčšiu teplotu ako 0 K [3].

## III. MODELOVANIE A ROZLOŽENIE TEPLOTNÝCH POLÍ V OKOLÍ ELEKTROENERGETICKÝCH ZARIADENÍ

Pre modelovanie a rozloženie teplotného poľa bol použitý softvérový nástroj ANSYS. ANSYS je simulačný softvérový balík, ktorý pri výpočtoch využíva metódu konečných prvkov (MKP). Metóda zabezpečuje získanie pomerne presného výsledku riešenia úloh. Úloha sa rozdelí na množstvo malých častí s konečným počtom, podľa ktorých sa následne zostavia algebrické rovnice.

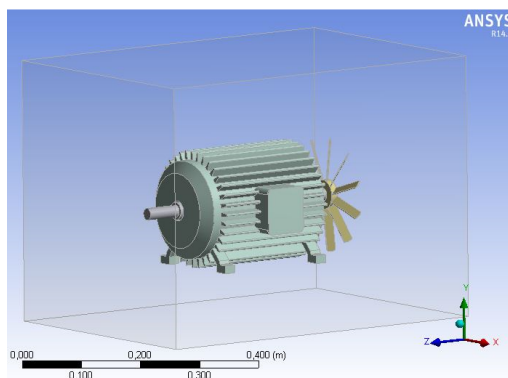
### A. Modelovanie a analýza 3-fázového asynchrónneho motora

Rozmery motora boli zistené odmeraním jednotlivých častí reálneho asynchrónneho elektromotora s označením Sg90L-4, ktorý bol pre tieto potreby demontovaný. Modelovanie sme začali od hriadeľa rotora, na ktorom sú rotorové plechy s drážkami pre vinutie rotorovej kľetky. Ďalej sme postupovali namodelovaním statora s počtom drážok  $Q_{m2} = 36$ . Jarmo statora je vyrobené zo statorových plechov, v ktorých je uložené vinutie. Vinutie od statorových plechov oddeľuje izolácia. Statorové vinutie je, na rozdiel od vinutia rotora, kde sú použité tyče, vyhotovené z drôtov. V simulácii je statorové vinutie kvôli zjednodušeniu vyhotovené ako tyče, taktiež v záujme zníženia náročnosti výpočtu nie sú jednotlivé póly fáz prepojené (Obr. 1).



Obr. 1 Telo rotora a statora a) bez vinutí (vľavo), b) s vinutiami (vpravo)

Po namodelovaní všetkých častí motora vrátane obalu bol simulovaný asynchrónny motor pomocou funkcie „enclosure“ umiestnený vo vzduchovom obale tvaru kvádra. V tomto priestore prebiehala simulácia (Obr. 2).



Obr. 2 Motor vo vzduchovom obale

Po dokončení úprav geometrie program vytvoril sieť (mesh) z elementov, ktoré slúžia k zostaveniu rovníc. V celom priestore „enclosure“ je nastavená hodnota konvekcie na hodnotu  $2 \text{ W} \cdot \text{m}^{-2} \text{ } ^\circ\text{C}$ . Táto podmienka zabezpečuje, že povrch každého telesa vyžiarí do vzduchu energiu



2 W z jednotky plochy pri rozdielne teplôt 1 °C. Takto bolo zaručené rovnomernejšie rozloženie teplotného poľa a interakcia medzi plynom a pevnou látkou.

Celkové straty sú súčtom mechanických strát, strát v železe, strát vo vinutí statora a v kľietke rotora a mechanických strát.

$$\Delta P = \Delta P_{Fe} + \Delta P_{j1} + \Delta P_{j2} + \Delta P_{mech} + \Delta P_d = 63 + 92,473 + 107,249 + 12,3 + 7,5 = 282,486 \text{ W} \quad (1)$$

Aby mohli byť zadané hodnoty strát do statorového vinutia, rotorového vinutia, statorových a rotorových plechov, je potrebné ich prepočítať na objem tak, aby ich bolo možné zadávať v simulácii do jednotlivých prvkov transformátora pomocou voľby „Internal Heat Generation“.

Energia vyžiarená zo železa statora:

$$q_{Fe,stat} = \frac{\Delta P_{Feh}}{V_{stat}} = \frac{57,25}{8,6836 \cdot 10^{-4}} = 65928877 \text{ W} \cdot \text{m}^{-3} \quad (2)$$

takto vypočítaný tepelný vnútorný zdroj bude zadaný na statorové plechy.

Energia vyžiarená z vinutia statora:

$$q_{j1,stat} = \frac{\Delta P_{j1}}{V_{stat}} = \frac{92,437}{4,035 \cdot 10^{-4}} = 229087980 \text{ W} \cdot \text{m}^{-3} \quad (3)$$

takto vypočítaný tepelný vnútorný zdroj bude zadaný na vinutie statora.

Energia vyžiarená zo plechov (železa) rotora:

$$q_{Fe,rot} = \frac{\Delta P_{mech} + \Delta P_d + \Delta P_v}{V_{rot}} = \frac{12,3 + 7,5 + 5,725}{5,1993 \cdot 10^{-4}} = 49093147 \text{ W} \cdot \text{m}^{-3} \quad (4)$$

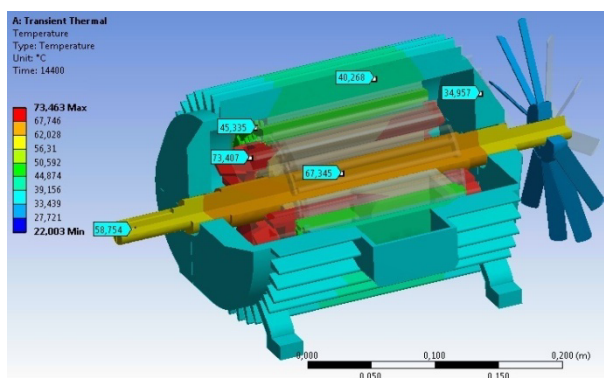
takto vypočítaný tepelný vnútorný zdroj bude zadaný na rotorové plechy.

Energia vyžiarená z kľietky rotora:

$$q_{j2,rot} = \frac{\Delta P_{j2}}{V_{rot}} = \frac{107,249}{3,5399 \cdot 10^{-4}} = 302971835 \text{ W} \cdot \text{m}^{-3} \quad (5)$$

takto vypočítaný tepelný vnútorný zdroj bude zadaný na vinutie rotora.

Teplota okolia v tejto simulácii bola nastavená na hodnotu  $\vartheta_0 = 22 \text{ °C}$ , čo predstavuje bežnú teplotu pri interiérových aplikáciách asynchrónneho motora.



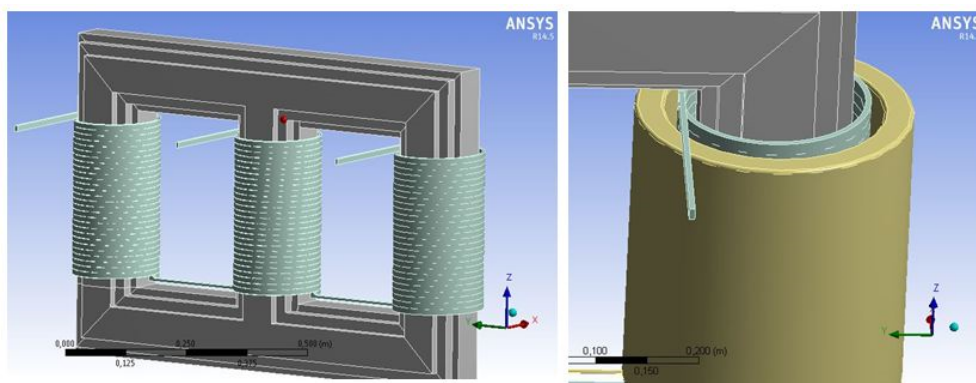
Obr. 3 Rozloženie teplotného poľa motora

Obsluha motora prichádza do kontaktu prevažne s obalom asynchrónneho motora, na povrchu ktorého v prevádzke pri nominálnom zaťažení je teplota podľa simulácie 40,3 °C, čo je iba o niečo vyššia teplota, ako teplota ľudského tela. Takýto dotyk nepredstavuje nebezpečenstvo úrazu, nehrozia popáleniny. Maximálna nameraná teplota bola  $\vartheta_{max1} = 73,463 \text{ °C}$ , a to vo vodičoch kľietky rotora, maximálne oteplenie nepresiahlo dovolené oteplenie  $\Delta \vartheta_{dovF} = 105 \text{ °C}$ .

Podmienka dovoleného oteplenia bola teda dodržaná. Chladenie je dostačujúce, aj keď nie je zabezpečený nútený obeh vzduchu. Avšak, udržiavanie vlastného chladenia v prevádzke, t.j. kontrola ventilátora, chladiacich mriežok a čistenie nánosov na chladiacich rebrách priaznivo vplyva na predĺženie životnosti elektrického stroja.

#### B. Modelovanie a analýza suchého distribučného transformátora 22/0,4 kV

Pri modeli transformátora bolo vytvorené najprv železné jadro. Kvôli optimalizácii výpočtu nebolo jadro vytvorené z jednotlivých plechov, ale z jedného kusu kovu. Ďalej nasledovalo sekundárne vinutie navinuté na jadre, ktoré je z profilového vodiča obdĺžnikového prierezu. Primárne vinutie je z vodiča kruhového prierezu, avšak nebolo možné vytvoriť presný model primárneho vinutia so skutočným usporiadaním závitov vinutia kvôli nedostatočnému výkonu počítača (pracovnej stanice).



Obr. 4 Modelovanie transformátora a) sekundárne vinutie (vľavo) b) primárne vinutie (vpravo)

V snahe dosiahnuť výsledky čo najpresnejšie, bolo zjednodušené primárne vinutie. Hodnota objemu, ktorú navinutý vodič zaberá, bola len minimálne rozdielna od objemu, ktorý zaberá primárne vinutie v simulácii.

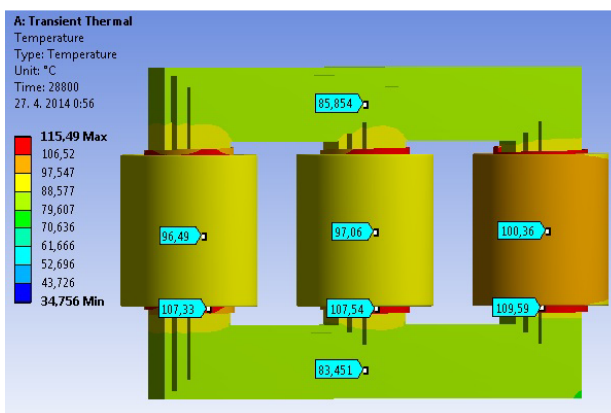
Podobne, ako pri asynchrónnom motore, bol namodelovaný transformátor umiestnený pomocou funkcie „enclosure“ do vzduchového obalu tvaru kvádra. Tento obal reprezentuje skriňu, kde bol umiestnený transformátor. Jej rozmery sú 1130×1470×830 mm. V priestore, ohraničenom obalom (enclosure) prebiehala simulácia. Hraničné a okrajové podmienky boli nastavené podobne, ako pri simulácii asynchrónneho motora.

Celkové straty predstavujú súčet strát (3-f) v primárnej cievke, sekundárnej cievke a strát v železnom jadre.

$$\Delta P_C = 3 \cdot \Delta P_{1Al} + 3 \cdot \Delta P_{2Al} + \Delta P_{Fe} = 40683 + 36288 + 12211 = 89182 \text{ W} \quad (6)$$

Vypočítané straty je potrebné prepočítať na objem, aby ich bolo možné zadať pre jednotlivé prvky transformátora pomocou voľby „Internal Heat Generation“ (podobne, ako v predchádzajúcom prípade, pri asynchrónnom motore).

Teplota okolia v tejto simulácii bola nastavená na hodnotu  $\vartheta_0 = 22 \text{ }^\circ\text{C}$ .



Obr. 5 Transformátor 22/0,4 kV

Maximálna nameraná teplota bola  $\vartheta_{\max 1} = 115,49 \text{ }^\circ\text{C}$ , na sekundárnom vinutí. Dôvodom je nízka hodnota prúdiaceho vzduchu a tiež veľká hodnota odoberaného prúdu. Podmienka dodržania hranice maximálneho oteplenia bola splnená. Chladenie je dostačujúce, avšak ak stúpne okolitá teplota nad normou stanovenú maximálnu teplotu okolia  $\vartheta_{40} = 40 \text{ }^\circ\text{C}$ , teplota transformátora bude  $\vartheta_{\max 1} = 135,49 \text{ }^\circ\text{C}$ , čo sa už blíži maximálnej dovolenej teplote pre izoláciu triedy F, t.j.  $\vartheta_{\text{dovF}} = 155 \text{ }^\circ\text{C}$ . Odporúčame preto aplikovať aktívne chladenie transformátora, čo predĺži životnosť transformátora a zníži straty.

### Návrh chladenia

Pre tento prípad predpokladajme, že počas 1 hodiny prevádzky transformátora, bude generované teplo transformátora rovné  $Q_{\text{strat}}$ .

$$Q_{\text{strat}} = \Delta P_C \cdot t = 89182 \cdot 3600 = 321,055 \text{ MW} \cdot \text{s} = 321,055 \text{ MJ} \quad (7)$$

kde  $t = 3600 \text{ s}$ .

Ak zanedbáme teplo, ktoré je potrebné na ohriatie vinutí a železného jadra a budeme predpokladať, že všetko teplo, ktoré vznikne prechodom prúdu bude teplo, ktoré ohrieva vzduch, môžeme určiť objem vzduchu, ktorý bude týmto teplom ohriaty o požadované

oteplenie  $\Delta\vartheta_{\text{pož}}$ . Uvažujme s oteplením z hodnoty 22 °C na 85 °C, čo predstavuje oteplenie  $\Delta\vartheta = 63$  °C.

Potom zo vzťahu pre výpočet tepla  $Q$ :

$$Q = c \cdot m_{\text{vzd}} \cdot \Delta\vartheta \quad (8)$$

kde  $c_{\text{vzd}} = 1003 \text{ J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$  je merná tepelná kapacita vzduchu,  $m_{\text{vzd}}$  je hmotnosť vzduchu,

Objem vzduchu  $V_{\text{vzd}}$  vyjadríme zo vzťahu pre hmotnosť vzduchu  $m_{\text{vzd}} = \rho_{\text{vzd}} \cdot V$ , a následne aký objem vzduchu sa ohreje za hodinu:

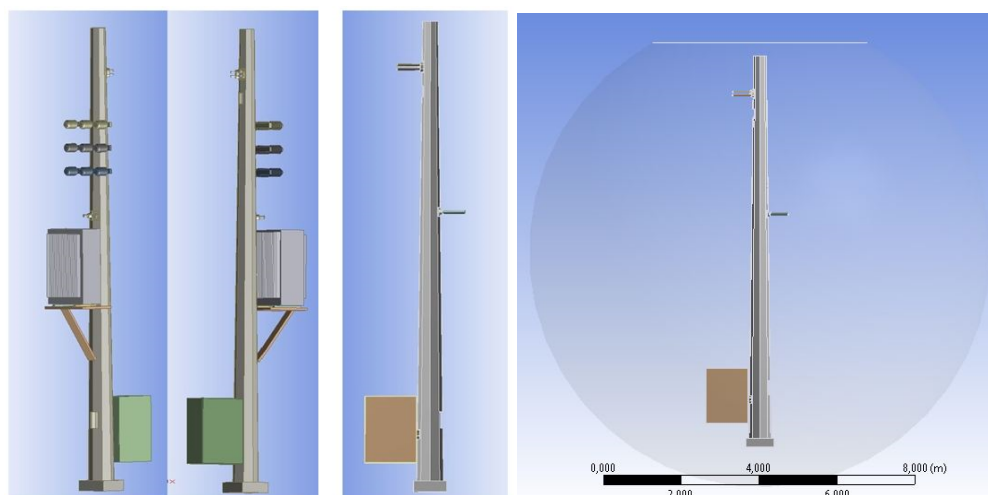
$$V_{\text{vzd}} = \frac{Q}{c \cdot \Delta\vartheta \cdot \rho_{\text{vzd}}} = \frac{8918,2 \cdot 3600}{1,003 \cdot 10^3 \cdot 63 \cdot 1,2047} = 421,754 \text{ m}^3 \quad (9)$$

kde  $\rho_{\text{vzd}} = 1,2047 \text{ kg} \cdot \text{m}^{-3}$  je objemová hustota vzduchu pri teplote 20 °C.

Teplu  $Q_{\text{strat}}$  ohreje vzduch o objeme  $V_{\text{vzd}} = 421,754 \text{ m}^3$  z teploty  $\vartheta_1 = 22$  °C na  $\vartheta_2 = 85$  °C. Keďže sme predpokladali, že straty  $\Delta P_c$  sa vyžiaria ako teplo za jednu hodinu, môžeme povedať, že na udržanie konštantnej hodnoty oteplenia pri normálnej prevádzke potrebujeme zabezpečiť, aby sa v okolí transformátora ohrial vzduch o objeme minimálne  $422 \text{ m}^3$  za hodinu. Takto dostávame hodnotu prietoku vzduchu  $422 \text{ m}^3 \cdot \text{h}^{-1}$ , resp.  $0,117 \text{ m}^3 \cdot \text{s}^{-1}$  prechádzajúci ventilačným zariadením, ktoré odvedie ohriaty vzduch zo skrine transformátora a zabezpečí požadovaný prísun chladného vzduchu z okolia na chladenie.

### C. Modelovanie a analýza ohraňovanej transformačnej stanice 22/0,4 kV 400 kW

Podstava modelu transformačnej stanice je tvorená 8-uholníkom. S rastúcou výškou sa prierez znižuje. Driek stožiaru je z konštrukčnej ocele.



Obr. 6 Modelovanie ohraňovanej trafostanice a) zjednodušený model (vľavo), b) vo vzduchovom obale (vpravo)

Kvôli náročnosti výpočtu, boli pre zjednodušenie geometrie odstránené prvky stožiarovej rozvodne, ktoré neovplyvňovali teplotné pole vo vnútri stožiaru. Taktiež boli odstránené ochranné zvodové rúrky, ktoré sa nachádzajú v drieku stožiaru. Transformačná stanica bola umiestnená do vzduchového obalu tvaru gule (enclosure), ktorý bol navrchu zrezaný.

Celkové straty transformačnej stanice pozostávali zo strát vo zvodových a vývodových vodičoch pre vzdušné vedenie:

$$\Delta P_C = 2 \cdot \Delta P_{\text{CYKY}} + 4 \cdot \Delta P_{\text{NAYY}} = 668,122 \text{ W} \quad (10)$$

Vypočítané straty je potrebné prepočítať na objem, aby ich bolo možné zadať pre jednotlivé prvky transformátora pomocou voľby „Internal Heat Generation“.

Energia vyžiarená z vodiča 1-CYKY-J 4×120:

$$q_{\text{CYKY}} = \frac{\Delta P_{\text{CYKY}}}{V_{\text{CYKY}}} = \frac{109,991}{3 \cdot 240 \cdot 10^{-6} \cdot 5,946} = 25688652 \text{ W} \cdot \text{m}^{-3} \quad (11)$$

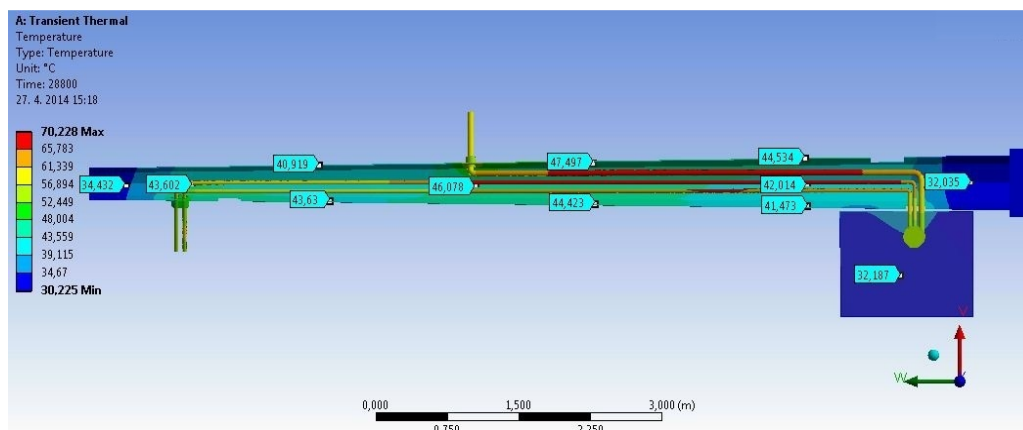
takto vypočítaný tepelný vnútorný zdroj bude zadaný pre vodič  $\text{CYKY}_1$  a  $\text{CYKY}_2$ .

Energia vyžiarená z vodiča NAYY-J 4×120:

$$q_{\text{NAYY}} = \frac{\Delta P_{\text{NAYY}}}{V_{\text{NAYY}}} = \frac{112,035}{3 \cdot 120 \cdot 10^{-6} \cdot 8,785} = 35425,773 \text{ W} \cdot \text{m}^{-3} \quad (12)$$

takto vypočítaný tepelný vnútorný zdroj bol zadaný pre vodič  $\text{NAYY}_1$ ,  $\text{NAYY}_2$ ,  $\text{NAYY}_3$  a  $\text{NAYY}_4$ .

Teplota okolia v tejto simulácii bola nastavená na hodnotu  $t_0 = 30\text{ }^{\circ}\text{C}$ , čo predstavuje teplotu počas letných dní.



Obr. 7 Transformačná stanica v reze s teplotami kovovej konštrukcie

Teplota vodičov bola podľa predpokladov najvyššia v dolnej tretine výšky stožiara. Pri uvažovanom usporiadaní, sa v danej časti nachádza šesť vodičov, ktoré sú priamo ohrievané prenášaným prúdom a nepriamo ohrievané jeden od druhého. Teplota jadier vodičov NAYY-J, bola  $t_{\text{NAYY}} = 70,2\text{ }^{\circ}\text{C}$ . Vodič CYKY-J mal teplotu  $t_{\text{CYKY}} = 67,8\text{ }^{\circ}\text{C}$ . Pracovná teplota udávaná výrobcom pre tieto vodiče je  $70\text{ }^{\circ}\text{C}$ , v simulácii bola táto teplota o čosi vyššia, avšak stále v súlade s dovoleným oteplením.

#### IV. ZÁVER

Tento príspevok sa venoval riešeniu rozloženia teplotného poľa v okolí vybraných elektroenergetických zariadení. Pri analýze teplotného poľa bolo prihliadané na to, aby nedochádzalo k prehrievaniu a tým k degradácii elektrických ako aj tepelnoizolačných vlastností jednotlivých častí zariadení. Následkom zvýšenej teploty môže byť deštrukcia zariadenia. Preto je dôležitý výber správnej izolácie s príslušnou teplotnou triedou, dimenzovanou na určené dovolené oteplenie. Samozrejme, že dovolené oteplenie a dovolená teplota nesmú prekročiť medzné hodnoty. Akékoľvek prekročenie tejto hodnoty znižuje životnosť izolácie a tým aj samotného prístroja. Každých  $10\text{ }^{\circ}\text{C}$ , nad dovolenú teplotu, znižuje životnosť stroja o polovicu. Z tohto vyplýva, že znižovanie teploty zvyšuje životnosť samotných zariadení. Preto je potrebné neustále snímanie teplôt, oteplenia a ich následné redukovanie, napríklad úpravami v konštrukcii alebo dodatočným chladením.

#### POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied grantom VEGA č. 1/0388/13.

#### REFERENCES

- [1] ANTAL, Š.: Termodynamika. Bratislava: STU, 2009. 255 s. ISBN 978-80-227-3212-3.
- [2] NOVÁK, P – KOLESÁR, J.: Elektrotepná technika. Bratislava: Alfa, 1990. 211 s. ISBN 80-05-00426-5.
- [3] KUMIČÁK, J.: Prenos tepla a hmoty. Košice: Elfa, 1994. 49 s. ISBN 80-88786-003-7.
- [4] HAVELSKÝ, V.: Chladiaca technika. Bratislava: STU, 2006. 168 s. ISBN 978-80-277-3212-3.
- [5] Forced Laminat Flow Over an Isothermal Plate. [online]. [citované 2014-04-20]. Dostupné na internete: <[https://www.efunda.com/formulae/heat\\_transfer/convection\\_forced/calc\\_lamflow\\_isothermalplate.cfm](https://www.efunda.com/formulae/heat_transfer/convection_forced/calc_lamflow_isothermalplate.cfm)>

# Nonlinear Effects in Optical Fiber and Wavelength Division Multiplexing

<sup>1</sup>Ján TÓTH, <sup>1</sup>Luboš OVSEŇÍK, <sup>1</sup>Ján TURÁN

<sup>1</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>jan.toth@tuke.sk, <sup>1</sup>lubos.ovsenik@tuke.sk, <sup>1</sup>jan.turan@tuke.sk

**Abstract** — Higher requirements on network broadband and capacities require some new approaches in future network design architectures. At the beginning we will provide a short overview about the single and multi-mode optical fibers as well as the condition for multimode regime determined by dimensionless parameter. Increasing of data-bit rate inevitably encounters natural difficulties in its realization. Rising bit-rates and the total network capacities cause nonlinear effects amplification. There are several nonlinear effect which have the most significant contribution to a signal deterioration and a total link attenuation. In this paper we will mostly pay attention to well-known phenomena's as Stimulated Raman Scattering, Stimulated Brillouin Scattering, Self-Phase Modulation, Cross-Phase Modulation and Four-Wave Mixing which appear in optical fiber. Wavelength division multiplexing allows us to use the fiber capacity with a high efficiency. This technology imposes the fact that we can use more carrier frequencies to maintain several informational streams within the fiber. This paper gives us an idea how the nonlinear effects are related to WDM systems and their importance in the concept of optical networks.

**Keywords** — Brillouin Scattering, Four-Wave Mixing, nonlinear effects, Raman Scattering, Self-Phase Modulation

## I. INTRODUCTION

The *Single mode optical* fiber usually consists of 4 layers. Right in the middle of the optical fiber lays an optical core. It is the thinnest part of this architecture. The diameter is usually around 5–8  $\mu\text{m}$ . This part basically leads the light beam through the optical fiber itself. The second part is called cladding and its proportions are usually 125  $\mu\text{m}$ . The third layer is buffer with 250  $\mu\text{m}$  in diameter (sometimes we can find a different name for this layer – a coating). All of these three layers are finally covered by the fourth part which is called jacket or buffer as well. A jacket is approximately 400  $\mu\text{m}$  thick. A single-mode fiber always propagates only one mode of light [1]. A quality of transmitted information is very high (in comparison to multi-mode fiber) and depends on the distance. It is often used for long distance network architectures and implementations.  $\text{GeO}_2$  and  $\text{P}_2\text{O}_2$  are used for core fabrication because they increase refractive index. On the other hand the B (boron) and F (fluorine) are accurate for cladding – decrease refractive index. The dimensionless parameter  $V$  defines the leading/guiding properties within the optical fiber. The equation (1) shows us that this parameter is dependent on several factors [1].

$$V = a(\omega/c)(n_1^2 - n_2^2)^{1/2}. \quad (1)$$

A parameter  $a$  represents a core diameter (radius),  $\omega$  is the light frequency and refractive indexes of core ( $n_1$ ) and cladding ( $n_2$ ). The dimensionless parameter  $V$  is important because it tells us how many light modes are we able to carry within optical fiber. If the value  $V$  is lower than 2.405 then the fiber is considered as single-mode fiber. A common core size for such optical fiber is less than 5  $\mu\text{m}$ . The optical fiber is also described by an effective refractive index which plays a significant role regarding to nonlinear effects in optical fiber [2].

The *Multimode optical fiber* is able to lead an optical signal with different modes (it is able to guide multiple modes). On the other hand the signal has much lower quality in comparison to single mode fiber. There are usually two values which define such kind of optical fiber and these are a core diameter and out cover diameter. Typical values for core are within a following range

50 – 62.5  $\mu\text{m}$ . Multi-mode fiber supports hundreds of modes. Multimode implementation is mostly used for shorter distances about hundreds of meters because of signal losses and its signal degradation. Passive optical networks (PON) are network architectures which deliver optical signal to the end users. For better understanding we can clarify that passive optical networks do not require any electrical power source [1,2]. There are different kinds of such PONs: FTTC (fiber-to-the-curb), FTTB (fiber-to-the-building) or FTTH (fiber-to-the-home). A PON usually consists of Optical Line Termination (OLT) and several Optical Network Units (ONUs) which are pretty much close to the end users.

## II. NONLINEAR EFFECTS IN OPTICAL FIBER

To describe an optical transmission process in optical fibers we have to consider nonlinearities which appear in optical fiber when the wave is propagating from receiver to the optical end.

We have to take into account a fact that any kind of dielectric material responds to light in a nonlinear way for intense electromagnetic field. As we know the optical fiber also belongs among dielectric elements. The base of nonlinear response has a lot of common with inharmonic motion of bound electrons in applied field conditions. The equation (2) which describes polarization will tell us more about nonlinear relations in light propagation process. The actual components represent the most significant nonlinear effects within an optical fiber [3].

$$P = \epsilon_0 \left( \chi^{(1)} \cdot E + \chi^{(2)} : EE + \chi^{(3)} : EEE + \dots \right) \quad (2)$$

In our consideration we will pay an attention to third-order susceptibility, which is responsible for third-harmonic generation, four-wave mixing and nonlinear refraction. It's important to say that most of the nonlinear effects come up from nonlinear refraction. The equation (3) shows us the simplest and the most general form of a refractive index calculation.

$$\tilde{n}(\omega, I) = n(I) + n_2 I = n + \bar{n}_2 |E|^2, \quad (3)$$

where  $n_2$  is an actual nonlinear coefficient related to third-order susceptibility from equation (2).

$$\bar{n}_2 = \frac{3}{8n} \text{Re}(\chi_{xxxx}^3). \quad (4)$$

### A. Refractive index based nonlinear effects

At this point we can define two main kinds of nonlinear phenomena's. The first group is represented by so called *Kerr nonlinearities*. Kerr effect itself occurs when light is coming through the crystal or glasses and its origin is a nonlinear polarization generated in medium. Self-phase modulation, Four-wave mixing and Cross-phase effects particularly belong to this first group. Kerr nonlinearities are characterized by the fact that the phase wave velocity is dependent on wave intensity. These effects also deal with the refractive index changes in optical fiber which is caused by the electrical instability/perturbation as it's been already mentioned equations (3),(4). Kerr based effects depend on wavelength on contrary to scattering phenomena's which go more over the intensity. So the refractive index is related to these phenomena [3].

### B. Scattering nonlinearities (intensity based)

Scattering nonlinearities cover the second nonlinear effects group. Sometimes it's also called inelastic photon scattering. It's about the exchanging of energy between waves with different wavelengths. In this process a part of field energy is transferred to nonlinear medium. The first of them is called Raman and the second one Brillouin scattering. Both of these can have spontaneous and stimulated characters. For the purposes of light wave communications the stimulated scatterings take part. The most meaningful difference between SRS and SBS is that for single mode optical fiber the stimulated Brillouin scattering can have only backward character in comparison to SRS [2,3].



### C. Raman scattering

Raman scattering occurs in two variants. A spontaneous Raman scattering occurs as a result of interaction of molecular vibrations in the (glass) material and incident optical wave. Incident photons cause a stimulated transmission of electrons and new photons with higher or lower wavelength are created. The result of such interaction between a group of incident and “vibration” photons can have two forms. Stokes radiation or Stoke line are the terms which basically refer for newly generated photons (waves) with higher wavelength. On the other hand the anti-Stokes line gains lower photon wavelength then the incident wave. The second form is called stimulated Raman scattering. There is a special condition which has to be fulfilled to achieve stimulated scattering. Energy of incident optical signal should pass some minimal (critical) value which has been theoretically calculated for particular wavelengths [3,4]. The power value for a single mode fiber at 1550 nm wavelength should be around 1.7 W for a single channel transmission. To complete an evaluation of SRS, the initial rising of Stoke wave can be formed by a single equation (5).

$$\frac{dI_S}{dz} = g_R I_P I_S, \quad (5)$$

where  $I_S$  is Stoke intensity,  $I_P$  is the pump intensity (an incidental field intensity) and  $g_R$  is Raman gain coefficient. The Raman gain parameter is experimentally measured and its values are shown in figure (Fig. 1).

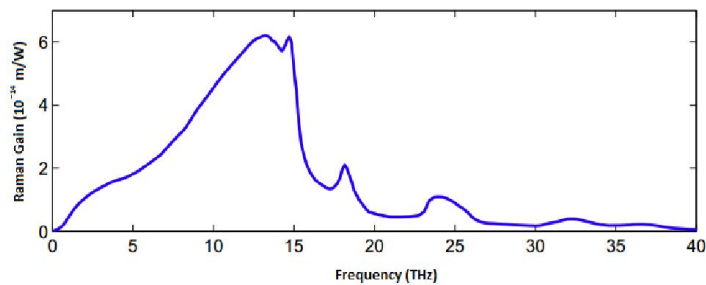


Fig. 1. The Raman gain coefficient curve [3]

The same formula is defined also for SBS and the only difference is that instead of  $g_R$  coefficient there is a subscript  $g_B$ . Both of these parameters were measured experimentally.

The stimulated Raman scattering is a process of frequency conversion and this feature is taken as an advantage when Raman lasers and amplifiers are designed. Such amplifiers have taken their place in light wave communication to amplify FDM (frequency division multiplexing) optical signals. Unfortunately there is one aspect in WDM and FDM optical systems when we are considering stimulated scattering. This scattering can cause transmission of energy (cross-talking) among multiple optical channels which has negative effects on particular signal channels (signal deterioration). That's why this phenomena has to be discussed when designing optical networks [3].

### D. Brillouin scattering

Brillouin scattering comes into view for significantly smaller incident optical signal power which is about 5-35 mW for single-channel transmission. It comes up by the interaction between incident optical wave and acoustic wave at the moments when wave vectors are matched. Brillouin scattering has only backward character in comparison to Raman scattering. The frequency changes (shifting) depend on wavelength of incident optical signal, a refractive index of the optical fiber core and the velocity of the acoustic waves. The frequency bandwidth of scattered wave (Brillouin signal) is considerably narrowed when comparing with Raman scattered signal bandwidth. This fact tells us that it could be a disadvantage for Brillouin amplifier implementation to low bit rate architectures. Brillouin scattering itself can raise attenuation in a WDM system (short waves channels).

### E. Four-wave mixing

Four-wave mixing or four-photon mixing are the synonyms for another nonlinear process in optical fibers. It is a special kind of Kerr effect. This phenomenon appears as the consequence of mixing an incident optical signal and signals which have been created because of Raman or

Brillouin stimulated scattering. Signals which have been formed after described mixing has feature both Stokes line and anti-Stokes waves. In general a four-photon mixing effect does not really have a positive effect on optical light wave multichannel architectures. Four-wave mixing can be also explained in other way. We have to realize that before launching any pump signal/light (specified by frequency) into the optical fiber there is already some amount of light (often called deteriorating signal/light). We usually have two pump signals which interact with probe signal which resulting to new signal forming with frequency. As shown in figure (see Fig. 2) the final equation for fourth wave frequency signal can be evolved as a product of frequencies summation in equation (6) [5].

$$f_{idler} = 2f_p - f_{probe} \quad (6)$$

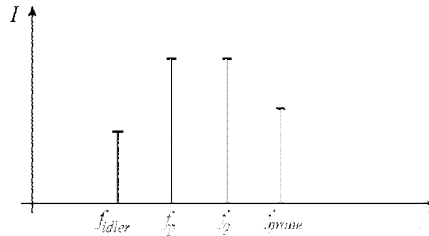


Fig. 2. Four wave mixing principle

#### F. Self-phase modulation

Self-phase modulation is the final nonlinear effect which is usually present when light propagates through the optical fiber. This effect causes that the spectrum of optical signal is spreading (a spectrum extending) during signal propagation. How much is spectrum spreading really depends on signal intensity. The frequency shifting depends on a refractive index. Phase magnitude can be extract from the following equation (7).

$$\phi = \tilde{n} k_0 L = \left( n + \bar{n}_2 |E|^2 \right) k_0 L, \quad (7)$$

where  $L$  is a fiber length and  $k_0$  is a wave number  $k_0 = 2\pi / \lambda$ . SPM is also responsible for soliton formatting in anomalous-dispersion fiber regime.

#### G. Cross-phase modulation

XPM is basically a phase shift of the optical field caused by presence of another field with a different wavelength/frequency, direction or polarization state. One can understand that as a total electric field equation (8).

$$E = 1/2 \hat{x} [E_1 \exp(-i\omega_1 t) + E_2 \exp(-i\omega_2 t) + \dots] \quad (8)$$

We can use equation (8) in case when two different electrical fields (frequencies  $\omega_1, \omega_2$ ) are polarized around x-axis and propagate simultaneously through the optical fiber. Than the nonlinear phase shift for the first field can be expressed as shown in equation (9).

$$\phi_{NL} = \bar{n}_2 k_0 L \left( |E_1|^2 + 2|E_2|^2 \right). \quad (9)$$

Cross-phase modulation has a significant contribution to a nonlinear phase shift when there are two optical fields with the same intensity and various wavelengths. In such a case a donation of XPM is twice of SPM.

### III. 3. WAVELENGTH DIVISION MULTIPLEXING

Research and understanding of nonlinear effects have a meaningful contribution when designing and implementing WDM technology to an existing network structures. Within the optical fiber there are numerous of nonlinear behaviors which have been already mentioned earlier. To increase the network/fiber capacity we can transmit several independent wavelengths (different colors) at the same time. The WDM system allows us to use an enormous fiber bandwidth. Many independent optical signals can be transmitted within one fiber [3].



The main factor which raises the nonlinearities in optical fiber is the intensity of electromagnetic field at the transmitter. The equation (10) gives us another way how to express a nonlinear phase shift.

$$\phi_{NL} = \bar{n}_2 k_0 L \left( |E_1|^2 + 2|E_2|^2 \right), \quad (10)$$

where  $\gamma$  is a nonlinear coefficient which depends on several factors from equation (11).

$$\gamma = \frac{n_2 \omega_0}{c A_{eff}} \quad (11)$$

$n_2$  is a cladding refractive index and  $A_{eff}$  is core cross-section (an effective area). It's important to note that the nonlinear phase shift is quite dynamic in fast moving optical pulse. In optical communications the maximum threshold of  $\phi_{NL}$  value for pulse has to be defined and it is regulated by the input optical power. It means that the phase shift should satisfy this condition  $\phi_{NL} < 1$ . So far  $\phi_{NL}$  was evaluated just for one channel. By considering WDM we are talking about multiple channels/frequencies and that means not only SPM has to be taken into account. The fact is that two and more channels will influence each other mostly in negative way. XPM results from different carrier frequencies of guided channels. Four-wave mixing has already been discussed. Mathematically the nonlinear phase shift  $\phi_{NL}$  for WDM architecture is given by equation (12) [8].

$$\phi_i^{NL} = \gamma \left[ \frac{1 - e^{-\alpha L}}{\alpha} \right] \left[ P_i + 2 \sum_{K=1, K \neq i}^w P_K \right] \quad (12)$$

Where  $w$  a total is number of channels and  $P_k$  is the power  $k$ th channel. The maximum phase shift is then represented by (13).

$$\phi_{max}^{NL} = \left[ \frac{\bar{\gamma}}{\alpha} \right] [2w - 1] P_i \quad (13)$$

SPM in optical communications doesn't act alone. GVD goes often side by side with a phase modulation over a length of the optical line. The power level a particular channel has to be patiently set up because the dispersion value has to be under some tolerable value (at a specific bit rate).

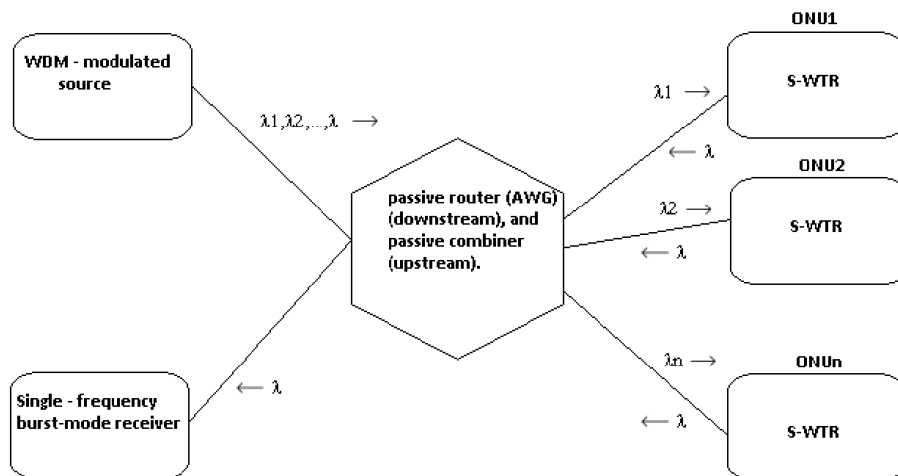


Fig. 3. A simple WDM-PON network architecture

WDM-PON networks is one the promising future architectures. A simple architecture is given by Figure (Fig. 3.), where S-WTR refers to Single-Wavelength transmitter and receiver [7].

#### IV. CONCLUSION

Optical nonlinearities in fibers give rise to a wealth of new effects. This is primarily due to the fact that, because these nonlinearities operate in a distributed manner, the effects they produce vary widely with the chromatic dispersion and birefringence characteristics of the fiber and, by consequence, with the wavelength, chirp, and polarization of the propagating light waves. In addition, several nonlinearities may act simultaneously, as in the examples of self-phase modulation (SPM) and cross-phase modulation (XPM), XPM and four-wave mixing (FWM), or stimulated Raman scattering (SRS) and FWM, resulting in an even greater variety of manifestations; super continuum generation (SCG) and soliton formation are two typical examples of combined nonlinearities. Nonlinear effects can be detrimental for optical communications, especially in wavelength division multiplexing (WDM) systems, where they can result in backscattering [stimulated Brillouin scattering (SBS)], noise (spontaneous Raman), pulse distortion (SPM, XPM, MI), and crosstalk between channels (XPM, FWM). Conversely, they are extremely useful for a variety of applications, from fiber lasers and amplifiers to wavelength converters, demultiplexers, optical switches, etc. Nonlinear effects will be particularly important in the next generation of optical networks, which will rely on all-optical functions for higher speed and greater capacity. All-optical functions should allow the partial elimination of optical-electronic-optical (O-E-O) conversion in optical networks, making them more transparent and more easily reconfigurable. The main challenge in all-optical networks will be controlling these nonlinearities and, in particular, their interplay. This will certainly require new types of fibers. For example, new glass compositions may provide intrinsically flatter SRS gain over a wider range of wavelengths, and new fiber designs, dispersion maps, and birefringence characteristics that can be precisely tuned to properly balance the desired nonlinearities.

#### ACKNOWLEDGMENT

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „063TUKE-4/2013 - The Use of Remote Controlled Optical Fibre Refractometer in Teaching“, and project no. „006TUKE-4/2014 - The Use of TUKE PON Experimental Model in Teaching. This work is also the result of the contract no. “APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks” supported by the Slovak Research and Development Agency.

#### REFERENCES

- [1] M. Cvijetic, “Coherent and Nonlinear Lightwave Communications,” Artech House, Inc. Norwood, MA 02062, 1996.
- [2] D. Mestdagh, “Fundamentals of Multiaccess Optical Fiber Networks,” Artech House, Inc. Norwood, MA 02062, 1995.
- [3] A. Govind, “Nonlinear Fiber Optics,” Elsevier, Inc. Waltham, MA 02451, Fifth edition 2013.
- [4] Y. Pochi, “Introduction to Photorefractive Nonlinear Optics,” John Wiley & Sons, Inc. New York, NY 10158-0012, 1993.
- [5] A. Osamu, T. Masateru, N. Shu, “Four-wave Mixing in Optical Fiber and Its Applications,” 1999.
- [6] B. Mukherjee, “WDM Optical Communication Networks: Progress and Challenges,” IEEE Journal on Selected Areas in Communications, Vol. 18, October 2000.
- [7] A. Banerjee, Y. Park, F. Clarke, H. Song, K. Kim, B. Mukherjee, “Wavelength-division-multiplexed Passive Optical Network (WDM-PON) Technologies for Broadband Access,” Journal of optical networking, 2005.
- [8] WDM network design. 4th chapter, 33 pages. [Online]. Available at: [http://media.techtarget.com/searchNetworking/Downloads/wdm\\_network\\_design.pdf](http://media.techtarget.com/searchNetworking/Downloads/wdm_network_design.pdf).

# Biometric Systems And Face Recognition

<sup>1</sup>Jana HANDRIKOVÁ, <sup>2</sup>Stanislav MARCHEVSKÝ

<sup>1</sup> Department of Informatics, Armed Forces Academy of General Milan Rastislav Štefánik, Liptovský Mikuláš, Slovak Republic

<sup>2</sup> Department of Electronics and Multimedia Communications, Technical University of Košice, Slovak Republic

<sup>1</sup>janka.handrikova@gmail.com, <sup>2</sup>stanislav.marchevsky@tuke.sk

**Abstract** — A number of applications tries to confirm identity of the persons requesting for the service by using their body characteristics. One of the most commonly used characteristics is the face recognition. This article shows main approaches used in face recognition systems.

**Keywords** — Biometric systems, Face recognition

## I. INTRODUCTION

Biometrics is being increasingly used today to establish person recognition in a large number of applications.

A number of biometric characteristics exist and are in use in various applications [1]: DNA, ear, face, facial, hand, and hand vein infrared thermogram, fingerprint, gait, hand and finger geometry, iris, keystroke, odor, palmprint, retinal scan, signature, voice. Each biometric has its strengths and weaknesses, and the choice depends on the application.

No single biometric is expected to effectively meet the requirements of all the applications.

Any human physiological and/or behavioral characteristic can be used as a biometric characteristic as long as it satisfies the following requirements [1]:

- Universality: each person should have the characteristic.
- Distinctiveness: any two persons should be sufficiently different in terms of the characteristic.
- Permanence: the characteristic should be sufficiently invariant (with respect to the matching criterion) over a period of time.
- Collectability: the characteristic can be measured quantitatively.

## II. STRUCTURE OF BIOMETRIC SYSTEMS

Depending on the application context, a biometric system may operate either in verification mode or identification mode.

In the verification mode, the system validates a person's identity by comparing the captured biometric data with her own biometric template(s) stored in the system database. In such a system, an individual who desires to be recognized claims an identity, usually via a personal identification number (PIN), a user name, or a smart card, and the system conducts a one-to-one comparison to determine whether the claim is true or not (e.g., "Does this biometric data belong to Bob?"). Identity verification is typically used for positive recognition, where the aim is to prevent multiple people from using the same identity [2].

In the identification mode, the system recognizes an individual by searching the templates of all the users in the database for a match. Therefore, the system conducts a one-to-many comparison to establish an individual's identity (or fails if the subject is not enrolled in the system database) without the subject having to claim an identity (e.g., "Whose biometric data is this?"). Identification is a critical component in negative recognition applications where the system establishes whether the person is who she (implicitly or explicitly) denies to be. The purpose of negative recognition is to prevent a single person from using multiple identities [2]. Identification may also be used in positive recognition for convenience (the user is not required to claim an identity). While traditional methods of personal recognition such as passwords, PINs, keys, and tokens may work for positive recognition, negative recognition can only be

established through biometrics [1].

The verification problem may be formally posed as follows: given an input feature vector  $X_Q$  (extracted from the biometric data) and a claimed identity  $I$ , determine if  $(I, X_Q)$  belongs to class  $w1$  or  $w2$ , where  $w1$  indicates that the claim is true (a genuine user) and  $w2$  indicates that the claim is false (an impostor). Typically  $X_Q$  is matched against  $X_I$ , the biometric template of  $X_Q$  corresponding to user  $I$ , to determine its category.

$$(I, X_Q) \in \begin{cases} w1, & \text{if } S(X_Q, X_I) \geq t \\ w2, & \text{otherwise} \end{cases}$$

where  $S$  is the function that measures the similarity between feature vectors  $X_Q$  and  $X_I$ , and  $t$  is a predefined threshold. The value  $S(X_Q, X_I)$  is termed as a similarity or matching score between the biometric measurements of the user and the claimed identity. Therefore, every claimed identity is classified into  $w1$  or  $w2$  based on the variables  $X_Q, I, X_I$  and  $t$  and the function  $S$  [1].

The identification problem, on the other hand, may be stated as follows. Given an input feature vector  $X_Q$  determine the identity  $I_k, k \in \{1, 2, \dots, N, N + 1\}$ . Here identity  $I_1, I_2, \dots, I_N$  are the identities enrolled in the system and  $I_{N+1}$  indicates the reject case where no suitable identity can be determined for the user. Hence

$$X_Q \in \begin{cases} I_k & \text{if } \max_k \{S(X_Q, X_{I_k})\} \geq t, k = 1, 2, \dots, N \\ I_{N+1}, & \text{otherwise} \end{cases}$$

where  $X_{I_k}$  is the biometric template corresponding to identity  $I_k$ , and  $t$  is a predefined threshold [1].

A biometric system is designed using the following four main modules [1]

1. Sensor module, which captures the biometric data of an individual.
2. Feature extraction module, in which the acquired biometric data is processed to extract a set of salient or discriminatory features.
3. Matcher module, in which the features extracted during recognition are compared against the stored templates to generate matching scores.
4. System database module, which is used by the biometric system to store the biometric templates of the enrolled users. The enrollment module is responsible for enrolling individuals into the biometric system database. During the enrollment phase, the biometric characteristic of an individual is first scanned by a biometric reader to produce a digital representation of the characteristic.

### III. FACE RECOGNITION

As for face recognition is a nonintrusive method, and facial images are probably the most common biometric characteristic used by humans to make a personal recognition. The applications of facial recognition range from a static, controlled “mug-shot” verification to a dynamic, uncontrolled face identification in a cluttered background (e.g., airport).

There are numerous application areas in which face recognition can be exploited [3]: Security, Surveillance, General identity verification, Criminal justice systems, Image database investigations, “Smart Card” applications, Multi-media environments with adaptive human-computer interfaces, Video indexing, Witness face reconstruction.

In addition to these applications, the underlying techniques in the current face recognition technology have also been modified and used for related applications such as gender classification, expression recognition and facial feature recognition and tracking; each of these has its utility in various domains: for instance, expression recognition can be utilized in the field of medicine for intensive care monitoring while facial feature recognition and detection can be exploited for tracking a vehicle driver’s eyes and thus monitoring his fatigue, as well as for stress detection. Face recognition is also being used in conjunction with other biometrics such as speech, iris, fingerprint, ear and gait recognition in order to enhance the recognition performance of these methods [3].

As in the other biometric system, we can define two main problems in face recognition. In identification problems, the input to the system is an unknown face, and the system reports back the determined identity from a database of known individuals, whereas in verification problems, the system needs to confirm or reject the claimed identity of the input face [4].

We can use different criteria to divide methods of face recognition in some groups. For example dimension of the image (2D, 3D images) or the type of the image (Image-based algorithm and Video-based algorithm for face recognition).

#### A. Process of face recognition on still images

The process of face recognition we can divide in a few steps [5]:

1. Detection and localization of the face on the image.
2. Normalization.
3. Extraction of the attributes.
4. Classification.

The goal of face detection is to determine whether or not there are any faces in the image, and, if present, to return the face dimensions in the keyframe. Face detection is a challenging task, since several factors influence the appearance of the face in the image. These include identity, pose (frontal, half-profile, profile), presence or absence of facial features such as beards, moustaches and glasses, facial expression, occlusion and imaging conditions [6]. For instance, illumination changes caused by light sources at arbitrary positions and intensities contribute to a significant amount of variability [7].

In a detection problem, two statistics are important: true positives (also referred to as detection rate) and false positives (reported detections in nonface regions). An ideal system would have very high true positive and very low false positive rates. In practice, these two requirements are conflicting [4].

**Normalization:** Once the face has been detected (separated from its background), the face needs to be normalized. This means that the image must be standardized in terms of size, pose, illumination, etc. [5]

**Feature Extraction:** Once the face image has been normalized, the feature extraction and recognition of the face can take place. In feature extraction, a mathematical representation called a biometric template or biometric reference is generated, which is stored in the database and will form the basis of any recognition task [5].

**Face Recognition:** In last phase the biometric template of the suspect face is compared with the biometric template of each face present in the database. We recognize the face when we get a match between these two biometric templates [5].

The usage of a mixture of techniques makes it difficult to classify these systems based purely on what types of techniques they use for feature representation or classification. In [4] the following categorization is used:

1. Holistic matching methods. These methods use the whole face region as the raw input to a recognition system [4]. In these methods, each face image is represented as a single high dimensional vector by concatenating the gray values of all pixels in the face [8].
2. Feature-based (structural) matching methods. Typically, in these methods, local features such as the eyes, nose, and mouth are first extracted and their locations and local statistics (geometric and/or appearance) are fed into a structural classifier [4]. In local methods, the original face is represented by a set of low dimensional local feature vectors, rather than one single full high dimensional vector [8].
3. Hybrid methods. Just as the human perception system uses both local features and the whole face region to recognize a face, a machine recognition system should use both. One can argue that these methods could potentially offer the best of the two types of methods [3]. The key factors that influence the performance of hybrid methods include how to determine which features should be combined and how to combine [8].

The main advantage of the holistic approaches is that they do not destroy any of the information in the images by concentrating on only limited regions or points of interest. However, this same property is their greatest drawback, too, since most of these approaches start out with the basic assumption that all the pixels in the image are equally important. Consequently, these techniques are not only computationally expensive but require a high degree of correlation between the test and training images, and do not perform effectively under large variations in pose, scale and illumination, etc. Nevertheless, several of these algorithms have been modified and/or enhanced to compensate for such variations [3].

The main advantage offered by the featured-based techniques is that since the extraction of the feature points precedes the analysis done for matching the image to that of a known individual, such methods are relatively robust to position variations in the input image. In principle, feature-based schemes can be made invariant to size, orientation and/or lighting. Other benefits of these schemes include the compactness of representation of the face images and high speed matching. The major disadvantage of these approaches is the difficulty of automatic feature detection and the fact that the implementer of any of these techniques has to make arbitrary decisions about which features are important. After all, if the feature set lacks

discrimination ability, no amount of subsequent processing can compensate for that intrinsic deficiency [3].

Hybrid approaches use both holistic and local features. The example of this approach is the system described in [10] based on Flexible Appearance model.

### *B. Video-based Face Recognition*

The video-based recognition has superior advantages over the image-based recognition.

- Possibility of employing redundancy contained in the video sequence to improve still images recognition performance [8].
- Temporal information of faces can be utilized to facilitate the recognition task. For example, the person-specific dynamic characteristics can help the recognition [11].
- More effective representations, such as a 3D face model or super-resolution images, can be obtained from the video sequence and used to improve recognition results [11].
- Video-based recognition allows learning or updating the subject model over time. Current and past frames in and are used to update the subject models to improve recognition results for future frames [11].

There also exists some disadvantages [11]:

- poor video quality
- low image resolution
- other influence factor (such as illumination, pose change, motion, occlusion, decorations, expressions, large distance from camera)

Most of the existing systems address video-based face recognition problems as follows: First, detect face and track it over time. Sometimes selecting good frames which contain frontal faces or valued cues is necessary. Next, when a frame satisfying certain criteria (size, pose, illumination and etc.) is acquired, recognition is performed, sometimes, by using still-to-still recognition technique [8]. A video-based face recognition system typically consists of three modules: one for detecting the face; a second one for tracking it; and a third one for recognizing it [3].

Given a frame of a video sequence, the detection module locates face candidates, while the tracking module finds the exact position of facial features in the current frame based on an estimate of face or feature locations in the previous frame(s). The recognition module identifies or verifies the face, integrating information from previous frames [12].

Face detection segments the face areas from the background [12]. During detection process, lots of traditional methods for still images can be introduced [11].

In general, face tracking can be divided into [11]:

- head tracking - the head is viewed as a rigid object performing translations and rotations [12]
  - color-based
  - model-based
  - shape-based
- facial features tracking - facial features deformations due to facial expressions or speech are viewed as non rigid transformations limited by the head anatomy [12]
- combination of head and facial feature tracking

Approaches for video-based recognition [5]:

- Spatio-temporal information based approaches
- Statistic model based approaches
- Hybrid cues based approaches – These methods utilize other cues obtained from video sequences, such as voice, gait, motion etc.

Video-based face recognition algorithms can have three possible learning modes, namely offline batch learning, online learning and hybrid learning. In batch learning the classifier is trained offline once only. The system is not updated unless a new identity is to be added in the database. In online learning, the system is completely trained online, however manual labeling of identities is still required. The hybrid learning approach learns generic (or specific) face models offline in a batch mode and continuously updates them during online recognition [12].

Temporal continuity in video allows tracking of facial features, which can help in compensating pose or expression variations, while motion, gait and other features may enhance the performance of face recognition. Moreover, the simultaneous comprehensive exploitation of spatiotemporal cues results in increased tracking and identification accuracy [13].

#### IV. CONCLUSION

Biometric based systems are becoming more popular due to their non-invasive nature. In our research we would like to focus on the video detection and recognition face system because there is the possibility to increase system accuracy by using the faces across a wide range of variations in pose and lighting. A face can be seen as a composition of micro structures depending on local situation. One of the widely applied descriptor for face detection and recognition is Local Binary Patterns (LBP) because of its resistance to pose and illumination changes, low computational complexity and ability to code fine details.

#### REFERENCES

- [1] Anil K. Jain, Arun Ross, Salil Prabhakar: An Introduction to Biometric Recognition, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 14, NO. 1, JANUARY 2004
- [2] J. L. Wayman: Fundamentals of biometric authentication technologies, Int. J. Image Graphics, vol. 1, no. 1, pp. 93–113, 2001.
- [3] Jafri, R. and Arabnia, H. R.: A Survey of Face Recognition Techniques. Journal of Information Processing Systems, 2009, 5, 2, 41–68.
- [4] Zhao, W., Chellappa, R. Rosenfeld, A., Phillips, P.J.: Face Recognition: A Literature Survey, ACM Computing Surveys, 2003, pp. 399-458
- [5] Ambika Ramchandra, Ravindra Kumar: Overview Of Face Recognition System Challenges .INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 2, ISSUE 8, AUGUST 2013 ,ISSN 2277-8616
- [6] Stamou, G. - Kollias, S.: Multimedia content and the semantic web – Standard, Methods and tools, <http://onlinelibrary.wiley.com/doi/10.1002/0470012617.fmatter/pdf>
- [7] Nagaraju C., Srinu B., Srinivasa Rao E: An efficient Facial Features extraction Technique for Face Recognition system Using Local Binary Patterns .International Journal of Innovative Technology and Exploring Engineering (IJITEE) , ISSN: 2278-3075, Volume-2, Issue-6, May 2013\
- [8] Xiaoyang Tan, Songcan Chen, Zhi-Hua Zhou, Fuyan Zhang: Face Recognition from a Single Image per Person: A Survey, Pattern Recognition Journal, September 2006, pp. 1725-1745
- [9] Ayushi: A survey on Feature Extraction Techniques, International Journal of Computer Applications (0975-8887), Vol. 66 – No.11, March 2013
- [10] LANITIS, A., TAYLOR, C. J., AND COOTES, T. F. 1995. Automatic face identification system using flexible appearance models. Image Vis. Comput. 13, 393–401.
- [11] Xiaoming Liu, Tsuhan Chen.: Video-based Face recognition Using Adaptive Hidden Markov Model, Proceedings on the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition
- [12] Ajmal Mian: Unsupervised Learning from Local Features for Video-Based Face Recognition. In: 8th IEEE International Conference on Automatic Face and Gesture Recognition, 17 Sept - 19 Sept 2008, Amsterdam, The Netherlands. pp. 1-6. IEEE Computer Society Press. ISBN 978-1-4244-2154-1
- [13] Shailaja A Patil, Pramod J Deore: Video-based face recognition: a survey . World Journal of Science and Technology 2012, 2(4):136-139 ISSN: 2231 – 2587 Available Online: [www.worldjournalofscience.com](http://www.worldjournalofscience.com)

# Introduction To NoSQL Databases

<sup>1</sup>Jana HANDRIKOVÁ, <sup>2</sup>Stanislav MARCHEVSKÝ

<sup>1</sup> Department of Informatics, Armed Forces Academy of General Milan Rastislav Štefánik, Liptovský Mikuláš, Slovak Republic

<sup>2</sup> Department of Electronics and Multimedia Communications, Technical University of Košice, Slovak Republic

<sup>1</sup>janka.handrikova@gmail.com, <sup>2</sup>stanislav.marchevsky@tuke.sk

**Abstract** — NoSQL databases continues grow in relevance in 2012, in many cases serving as the data layer for Big Data application. This class of applications requires new functionality – ability to store and retrieve great quantities of poor structured or unstructured data too large for traditional relational databases on the different, geographic remote places. This article is a short introduction to NoSQL databases.

**Keywords** — Big Table, Document database, Graph database, Key value database, NoSQL database

## I. INTRODUCTION

In computing, NoSQL (commonly interpreted as "not only SQL") is a broad class of database management systems identified by non-adherence to the widely used relational database management system model. NoSQL databases are not built primarily on tables, and generally do not use SQL for data manipulation. NoSQL database systems are often highly optimized for retrieval and appending operations and often offer little functionality beyond record storage (e.g. key-value stores). The reduced run-time flexibility compared to full SQL systems is compensated by marked gains in scalability and performance for certain data models. In short, NoSQL database management systems are useful when working with a huge quantity of data when the data's nature does not require a relational model. The data can be structured, but NoSQL is used when what really matters is the ability to store and retrieve great quantities of data, not the relationships between the elements. Usage examples might be to store millions of key-value pairs in one or a few associative arrays or to store millions of data records [1].

Why NoSQL?

- Huge amount of unstructured and semi-structured data.
- The complexity of the data, we need more attributes to describe the entity.
- The data on the different, geographic remote places.
- Now, websites demand more interactive model.
- In the highly social applications typically using NoSQL on the back-end, consistency from a database standpoint is hard to achieve while still providing a responsive user experience. The distributed computing principle of "eventual consistency" describes how many NoSQL databases handle this issue in a parallel environment [2].

## II. CHARACTERISTICS OF NOSQL DATABASES

NoSQL systems work with data in memory, typically or uploading chunks of data from many disks in parallel. The issue is that "traditional" relational databases don't provide the same models and, thus, the same performance [8].

NoSQL database are schema-less / schema-free. Schema is typically dynamic - records can add new information on the fly, dissimilar data can be stored together as necessary [12]. The semantics of the data are embedded within a flexible connection topology and a corresponding storage model. This provides greater flexibility for managing large data sets while simultaneously reducing the dependence on the more formal database structure imposed by the relational database systems [14].

NoSQL databases use object-oriented APIs for **data manipulation**. UnQL – aspires to be the standard language for data manipulation for NoSQL databases. CouchBase uses Non-first Normal Form Query Language (N1QL) - a query language for document databases.

The distributed computing principle of "eventual consistency" (BASE attributes) describes



how many NoSQL databases handle this issue in a parallel environment. **BASE** (**B**asically **A**vailable, **S**oft state, **E**ventual consistency):

- **Basic Availability.** NoSQL databases spread data across many storage systems with a high degree of replication.
- **Soft State.** Data consistency is the developer's problem and should not be handled by the database.
- **Eventual Consistency.** The distributed computing principle of “eventual consistency” describes how many NoSQL databases handle this issue in a parallel environment.[2] Eventual consistency implies that there could be small intervals of inconsistency between replicated nodes as data gets updated among peer-to-peer nodes. Eventual consistency does not mean inconsistency. It just implies a weaker form of consistency than the typical ACID type consistency found in RDBMS [11]. Consistency depends on product. Some provide strong consistency (e.g., MongoDB) whereas others offer eventual consistency (e.g., Cassandra) [12].

**Scalability** is the ability of a system to increase throughput with addition of resources to address load increases [11]. Scaling horizontally, meaning that to add capacity, a database administrator can simply add more commodity servers or cloud instances. The database automatically spreads data across servers as necessary [12]. Processing data spread across a cluster of horizontally scaled machines is complex. The **MapReduce model** possibly provides one of the best possible methods to process large-scale data on a horizontal cluster of machines [11]. Map/Reduce is a multistep querying process, which takes a big task and breaks it down into multiple smaller tasks. The first step is to produce a Map which is executed on each server. Subsequent Reduce steps are executed on each individual server and then on groups of servers using the results of the previous Reduce step to come up with the final answer [13]. The steps of the Map/Reduce process are [14]:

1. Map/reduce framework will split the input data to each block.
2. The use of the map function.
3. This step is the combiner step.
4. This step is the reduce step.

### III. CATEGORIES OF NOSQL DATABASES

The NoSQL databases can be divided into four categories:

- Key Value
- Big Table (Column oriented)
- Document
- Graph DB

#### A. Key Value Databases

Key Value database is essentially a schema-less construct containing a key along with a piece of associated data or object [2].

Accessing memory locations in assembly-level programming is essentially following the Key-Value pattern, with the memory location's address serving as the key, and the value is stored at that memory address. The computer science concept of the hash table serves as another example, with a hash function transforming the key into an index used to find its associated value.

The reason this pattern gets used again and again in programming and computer engineering is because of speed. It can be argued that computers at their lowest (and fastest) level depend on the Key-Value pattern, so it makes perfect sense the Key-Value database is much more suitable for the requirements of massive Big Data when compared to the relational model. In short, Key-Value data stores are much faster [4].

The key value type basically, uses a hash table in which there exists a unique key and a pointer to a particular item of data. A bucket is a logical group of keys - but they don't physically group the data. There can be identical keys in different buckets.

Performance is enhanced to a great degree because of the cache mechanisms that accompany the mappings. To read a value you need to know both the key and the bucket because the real key is a hash (Bucket + Key) [7].

The key can be synthetic or auto-generated while the value can be String, JSON, BLOB (basic large object) [7].

The ACID (atomic, consistent, isolated, durable) principle normally present in relational databases is not supported by the massive throughput typical of many systems using a Key-Value database. Because of this, many Key-Value systems use an “eventually consistent” model, widely used in parallel processing and distributed systems [4].

It becomes quite clear that key value stores are great around the Availability and Partition aspects but definitely lack in Consistency [7].

Weaknesses of key value databases:

- One, is that the model will not provide any kind of traditional database capabilities (such as atomicity of transactions, or consistency when multiple transactions are executed simultaneously). Such capabilities must be provided by the application itself [7].
- Secondly, as the volume of data increases, maintaining unique values as keys may become more difficult; addressing this issue requires the introduction of some complexity in generating character strings that will remain unique among an extremely large set of keys [7].

An example of this type is **FoundationDB**.

**Data model:** FoundationDB’s core data model is an ordered key-value store. Also known as an ordered associative array, map, or dictionary, this is a common data structure composed of a collection of key-value pairs in which all keys are unique. Starting with this simple model, an application can create higher-level data models by mapping their elements to individual keys and values [22]. FoundationDB support a diverse set data models through simple layers which map different data models onto FoundationDB’s ordered keys and values.

**ACID and transactions:** All reads and writes in FoundationDB are accomplished using transactions. These transactions are fully ACID (Atomic, Consistent, Isolated, and Durable) and span multiple machines with high performance. FoundationDB’s isolation is the highest available - transactions appear to occur sequentially. FoundationDB’s durability is the strongest - all transactions are redundantly stored to disk before they are considered committed [23].

**Query language:** The FoundationDB core exposes a robust and powerful API but includes no separate query language. FoundationDB empowers developers to employ a broad range of data models and use the query languages best suited to their applications, implemented as layers [23].

**Indexes:** FoundationDB’s layer concepts lets us implement spatial indexing outside the database with a small amount of code, making it easy to inspect, extend or modify [21].

**Scalability:** FoundationDB offers scalability from partial utilization of a single core on a single machine to full utilization of dozens of powerful multicore machines in a cluster. A single FoundationDB instance can also be used by multiple tenants and host many virtual data stores. In this multiple-tenant mode, FoundationDB can scale to arbitrarily low loads, and pricing models such as metered pricing become possible [20].

### *B. Big Table (Column Store) databases*

**Data model:** Practically every other Big Table database in the industry (with some exceptions) is modeled after or inspired by Google’s Bigtable [5].

Big table is a distributed storage system for managing structured data that is designed to scale to a very large size: petabytes of data across thousands of commodity servers. Big table is a sparse, distributed, persistent multidimensional sorted map. The map is indexed by a row key, column key, and a timestamp; each value in the map is an uninterrupted array of bytes [5].

Data is stored in cells grouped in columns of data rather than as rows of data. Columns are logically grouped into column families. Column families can contain a virtually unlimited number of columns that can be created at runtime or the definition of the schema. Read and write is done using columns rather than rows [7]. Columnar databases store all the cells corresponding to a column as a continuous disk entry thus makes the search/access faster [7].

An example of this type is **Hbase**, an open-source, distributed, versioned, column-oriented store modeled after Google’s Bigtable [15].

**Data model:** A namespace is a logical grouping of tables analogous to a database in relation database systems. Tables are declared up front at schema definition time. Row keys are uninterpreted bytes. A {row, column, version} tuple exactly specifies a cell in HBase. Columns in Apache HBase are grouped into column families. All column members of a column family have the same prefix [17]. Table is made up of any number of regions. Region is specified by its startKey and endKey [18].

**Data manipulation:** The four primary data model operations are Get, Put, Scan, and Delete. Get returns attributes for a specified row. Put either adds new rows to a table (if the key is new) or can update existing rows (if the key already exists). Scan allow iteration over multiple rows for specified attributes. Delete removes a row from a table [16]. No built-in joins (use

MapReduce) [18].

**ACID:** Hbase is not ACID-compliant [19].

Atomicity

- All mutations are atomic within a row. Any put will either wholly succeed or wholly fail.
- APIs that mutate several rows will not be atomic across the multiple rows.
- The order of mutations is seen to happen in a well-defined order for each row, with no interleaving.

Consistency and Isolation

- All rows returned via any access API will consist of a complete row that existed at some point in the table's history.

Consistency of Scans

- A scan is not a consistent view of a table. Scans do not exhibit snapshot isolation.
- Those familiar with relational databases will recognize this isolation level as "read committed".

Durability

- All visible data is also durable data. That is to say, a read will never return data that has not been made durable on disk.
- Any operation that returns a "success" code (e.g. does not throw an exception) will be made durable.
- Any operation that returns a "failure" code will not be made durable (subject to the Atomicity guarantees above).
- All reasonable failure scenarios will not affect any of the listed ACID guarantees.

**Support of transactions:** There are no transactions [19].

**Indexes:** The row key is the primary key and has a backing index. Primary key could be a composite of values. HBase has no built-in support for secondary indexes [19].

**Scaling:** Horizontal scalability – just add hardware [19].

**Use:** Apache HBase provides Big table-like capabilities on top of Hadoop and HDFS.

### C. Document Databases

Data model is a collection of documents, a document is a key value collection. Document stores encompass a wide array of formats and binary encodings. Standard markup formats like XML and JSON combine with proprietary binaries like Microsoft Word and Adobe's PDF [2]. "IEEE").

Obviously, Document Stores center on the concept of data stored within a document. The encoding of this document can either be text in some form of a markup language or something formatted as binary. Examples of binary encoding include the aforementioned Microsoft Word or any other user created file managed by a proprietary software program, such as Excel, Photoshop, Pro Tools, and others [3].

Text markup languages used for data purposes are plentiful: from YAML to JSON to the nearly ubiquitous XML. Both flavors of document types feature in the overall world of Document databases [3].

Keys are normally used to retrieve items from a Document Store. Usually in a string format, the key can represent the path to a stored document, although performance improvements result from indexing those keys. In addition to faster access, indexes using other data types provide flexibility in the queries used for document retrieval – for example returning all stored documents containing a certain phrase [3].

An example of document database is **MongoDB**. On page <http://docs.mongodb.org/manual/> you can find the characteristics of MongoDB:

**Data model:**

- MongoDB database stores its data in collections. A collection holds one or more documents. Documents in a single collection may have a unique combination and set of fields. Document need not have identical fields. You can add a field to some documents in a collection without adding that field to all documents in the collection. The following table presents the various SQL terminology and concepts and the corresponding MongoDB terminology and concepts.
- MongoDB supports neither joins nor transactions.

However, it features secondary indexes, an expressive query language, atomic writes on per-document level, and fully-consistent reads.

**Support of SQL:** No. However, MongoDB does support a rich, ad-hoc query language of its own.

**Support of transactions:** MongoDB does not provide ACID transactions.

**Indexes:** Fundamentally, indexes in MongoDB are similar to indexes in other database

systems. MongoDB defines indexes at the collection level and supports indexes on any field or sub-field of the documents in a MongoDB collection.

**Scaling:** MongoDB features master-slave replication with automated fail over and built-in horizontal scaling via automated range-based partitioning.

**Typical uses for MongoDB:** Examples include content management systems, mobile applications, gaming, e-commerce, analytics, archiving, and logging.

#### D. Graph Databases

Graph databases use a matrix view of the underlying data, focused on the relationships between two entities [2]. The objects known as Nodes and Edges serve the roles of Entities and Relationships from standard SQL architecture. Nodes also contain properties which describe the actual data contained within each object. A diagram of a Graph database looks similar to the object diagrams used in object oriented programming [6].

The biggest advantage of Graph databases is obviously their speed for certain types of transactions; in particular, those involving relationships, since processing-intensive joins are not required. The fact that the design of Graph databases is less dependent on complex schema than their relational counterparts also lends itself to easier modifications and migrations to in-production systems [6].

An example of graph database is **Neo4j**.

**Data model:** A graph database stores data in a graph, the most generic of data structures, capable of elegantly representing any kind of data in a highly accessible way. The simplest possible graph is a single Node, a record that has named value referred to as Properties.

Topple the stacks of records in a relational database while keeping all the relationships, and you'll see a graph. Where an RDBMS is optimized for aggregated data, Neo4j is optimized for highly connected data.

**Support of SQL:** The traversal is the operation of visiting a set of nodes in the graph by moving between nodes connected with relationships. It's a fundamental operation for data retrieval in a graph, and as such, it is unique to the graph model. The key concept of traversals is that they're localized - querying the data using a traversal only takes into account the data that is required, without the need to perform expensive grouping operations on the entire data set, like you did with join operations on relational data [9].

**Index:** Some definitions require that a graph database use index-free adjacency, meaning that connected nodes physically "point" to each other in the database [10].

**Support of transactions:** Transactions in Neo4j are semantically identical to traditional database transactions. It should be pointed out that not all graph databases are fully ACID [9].

**Scaling:** The 1.9 release of Neo4j can support single graphs having tens of billions of nodes, relationships, and properties. This allows for graphs with a social networking dataset roughly the size of Facebook's [9].

**Typical use:** social applications, geospatial applications, master data management (identifying, cleaning, storing, and, most importantly, governing this data), network and Data Center Management.

## IV. CAP THEOREME

Brewer's Theorem was conjectured by Eric Brewer and presented by him as a keynote address at the ACM Symposium on the Principles of Distributed Computing (PODC) in 2000 [11].

**Consistency, Availability, and Partition Tolerance (CAP)** are the three pillars of Brewer's Theorem that underlies much of the recent generation of thinking around transactional integrity in large and scalable distributed systems. Succinctly put, Brewer's Theorem states that in systems that are distributed or scaled out it's impossible to achieve all three (Consistency, Availability, and Partition Tolerance) at the same time. You must make trade-offs and sacrifice at least one in favor of the other two. However, before the trade-offs are discussed, it's important to explore some more on what these three factors mean and imply [11].

**Consistency** is not a very well defined term but in the context of CAP it alludes to atomicity and isolation. Consistency means consistent reads and writes so that concurrent operations see the same valid and consistent data state, which at minimum means no stale data [11].

**Availability** means the system is available to serve at the time when it's needed [11].

**Partition Tolerance** - Parallel processing and scaling out are proven methods and are being adopted as the model for scalability and higher performance as opposed to scaling up and building massive super computers.

## V. CONCLUSION

Today, there is more than 150 NoSQL databases, the list of them you can find at

<http://nosql-database.org>.

#### REFERENCES

- [1] Praseeda Manoj :Emerging Database Models and Related Technologies . Volume 3, Issue 2, February 2013, ISSN: 2277 128X , International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Available online at: [www.ijarcsse.com](http://www.ijarcsse.com)
- [2] Paul Williams :The NoSQL Movement — What is it? ER World 2013, FREE on line Conference, Data Modeling at the Center of Data Management, November 12-14, 2013
- [3] Paul Williams :The NoSQL Movement: Document Databases ER World 2013, FREE on line Conference, Data Modeling at the Center of Data Management, November 12-14, 2013
- [4] Paul Williams :The NoSQL Movement: Key-Value Databases ER World 2013, FREE on line Conference, Data Modeling at the Center of Data Management, November 12-14, 2013
- [5] Paul Williams :The NoSQL Movement: Big Table Databases ER World 2013, FREE on line Conference, Data Modeling at the Center of Data Management, November 12-14, 2013
- [6] Paul Williams :The NoSQL Movement: Graph Databases ER World 2013, FREE on line Conference, Data Modeling at the Center of Data Management, November 12-14, 2013
- [7] Girish Kumar, Rahul Checker :Exploring the Different Types of NoSQL Databases
- [8] SQL and relational databases: They're not right for the cloud Cloud is causing many to leave SQL and relational database By David Linthicum | InfoWorld | Published: 11:31, 17 March 2010
- [9] Jonas Partner – Aleksa Vukotic- Nicky Watt:MEAP Edition Manning Early Access Program Neo4j in Action MEAP version ,2013 Manning Publications
- [10] Graph Databases by Ian Robinson, Jim Webber, and Emil Eifrem Copyright © 2013 Neo Technology, Inc.,ISBN: 978-1-449-35626-2
- [11] Shashank Tiwari : PROFESSIONAL NoSQL , Published by John Wiley & Sons, Inc. , 2011, ISBN: 978-0-470-94224-6
- [12] <http://www.mongodb.com/learn/nosql>
- [13] Charles Nurse: Look Mom, NoSQL! - 7. Map Reduce explained, [explained.aspx](http://explained.aspx)
- [14] David Loshin: An Introduction to NoSQL <http://data-informed.com/introduction-nosql-data-management-big-data/>
- [15] <http://hbase.apache.org/>
- [16] [http://hbase.apache.org/book/data\\_model\\_operations.html](http://hbase.apache.org/book/data_model_operations.html)
- [17] <http://hbase.apache.org/book.html#table>
- [18] <http://zhangjunhd.github.io/2013/02/25/apache-hbase/>
- [19] James Chin, Zikai Wang : Hbase . A Comprehensive Introduction (presentation) <http://cs.brown.edu/courses/cs227/archives/2011/slides/mar14-hbase.pdf>
- [20] <https://foundationdb.com/white-papers/scalability>
- [21] <https://foundationdb.com/blog/spatial-indexing-layer-for-foundationdb>
- [22] <https://foundationdb.com/documentation/data-modeling.html>
- [23] <https://foundationdb.com/features>

# Bagging Efficiency of Machine Learning in Steganalysis

<sup>1</sup>Vladimír HAJDUK, <sup>1</sup>Martin BRODA, <sup>1</sup>Dušan LEVICKÝ

<sup>1</sup> Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

vladimir.hajduk@student.tuke.sk, martin.broda@tuke.sk, dusan.levicky@tuke.sk

**Abstract** — This paper is focused on image steganalysis working in DCT domain. The mentioned method is known as feature based steganalysis and belongs to blind steganalysis. In blind steganalysis, trained model is not limited to detect secret message embedded only by one steganography algorithm used in training. Traditional techniques use individual classifiers to detect presence of secret message. The well known and common used individual algorithm is support vector machines (SVM). In this work comparing of efficiency of SVM with ensemble classifier using the method called bagging. Bagging is method of using combined learner consist of many weak learners, each trained on (uniformly) randomly selected subspace of the feature space. This mainly ensures shortly training time.

**Keywords** — bagging, features based steganalysis, ensemble learning, support vector machines

## I. INTRODUCTION

The goal of steganalysis is to detect secret message hidden in multimedia data, such as video, images and audio. Modern steganalysis has three different purposes: to detect, extract and erase hidden information. Each of these purposes may be achieved separately. Common methods are focused on detection secret message. Output of detection indicates whether the analyzed multimedia obtains embedded secret information or not. The second domain interested in extraction hidden information is called forensic steganalysis. The goal of forensic steganalysis is extraction content of secret message. Last method of steganalysis attack is erasing or damaging secret message obtained in multimedia data [1].

High computational complexity of modern techniques of steganalysis takes a long time to train quality model used to detect presence of steganography. The method of mentioned training is called machine learning. Machine learning is science discipline belongs to artificial intelligence. It is inspired by human learning system and gives this ability of self learning to machines, too. Machine learning is used to solve two main problems: classification and sequential problems. In classification learner makes a decision to classify some problem to one of classes. If these classes are presented in training, it is supervised learning. In addition, in sequential problems learner knows start and finish position and seeks road to achieve that. It is called unsupervised learning [2].

This paper discusses the image steganalysis, i.e. steganalysis detecting secret message hidden in and transmitted by static images.

## II. BLIND STEGANALYSIS

Due to depending on steganography algorithms two methods of image steganalysis are recognized: *target steganalysis* and *blind steganalysis*.

Target steganalysis depends on the used steganographic algorithm in training phase and detects presence of that steganographic algorithm only.

On the other hand, blind steganalysis gives sufficient results of detecting the secret message embedded by new, no well known algorithm even.

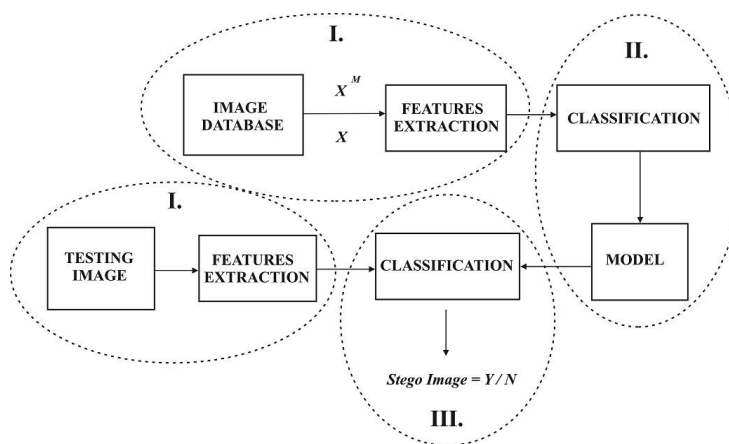


Fig. 1. Model of Blind Steganalysis. I. Phase of extraction statistical characteristics. II. Model training phase. III. Testing phase.

Model of blind steganalysis is shown in Fig. 1. The first phase (marked as I.) is the phase of extraction characteristic features from image. The images are from image database that consists of cover images (without secret message, marked as  $X$ ) and corresponding stego images (with embedded secret information,  $X^M$ ). Testing image is the image that we want to know whether is stego or cover. Features are extracted in next block with characteristic name. Phase II. is called training phase and works with features obtained in block mentioned before. Classifier marked extracted features vectors as stego or cover according class where they belong to. There is method of machine learning (more accurate supervised learning) used to train model based on input arguments. This model is used in training phase (III.) to determine class the testing image belong to (whether testing image is stego or not).

#### A. Feature Based Steganalysis

Image steganalysis works in spatial and frequency domains. It depends on domain where statistical features are obtained. Features from spatial domain are called Binary similarity measures (BSM). Steganalysis using that features detects decrease of correlation between neighboring bits and between bits of either bits plain after embedding secret information. In frequency domain there are both DWT (Discrete Wavelet Transformation) and DCT (Discrete Cosine Transformation) transformation used. Steganalysis of features from wavelet domain is known as Wavelet based steganalysis (WBS) and when features are extracted from DCT domain we discuss about Features based steganalysis (FBS).

FBS, in some letters called as Features calibrated Based Steganalysis, uses calibration to estimation of cover image (image without secret message).

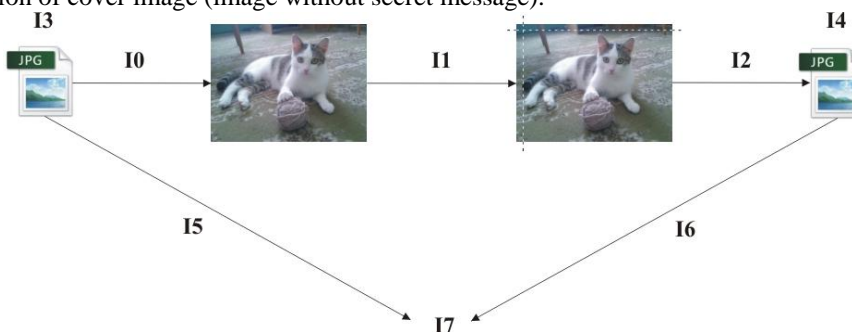


Fig.2. Estimation of cover image via calibration. Features extracted from image before calibration (left) and reference image (right), both in DCT domain.

In Fig. 1 the process of calibration is illustrated. Firstly, JPEG image (I3) is decompressed from DCT to spatial domain and cropped by 4 pixels in horizontal and vertical direction. Output image is called reference image and estimates cover image. No matter whether image I3 is stego (with hidden information) or cover. At this point features are extracted from both, image before calibration (I5) and reference image (I6) and result – distinction between them (I7) is vector of calibrated features. Calibrated features are inputs of training phase stegoanalysis system [3] [4].

In this paper we discuss features based steganalysis (FBS) that works in DCT domain.

### III. CLASSIFIERS

Classifiers in steganalysis are used in training and testing phases of blind steganalytic system. Classifiers are able to put testing object to the appropriate class. Among the most popular individual classifiers we can include [4]:

- Support Vector Machines (SVM)
- Fisher Linear Discriminate (FLD)
- Quadratic Discriminant Analysis (QDA)
- Ordinary Least-Squares regression (OLS)
- Bayesian Belief Networks (BBNs)
- Naive Bayes Classifiers (NBCs)

And the most accurate and the most common used are Support Vector Machines (SVM) and Fisher Linear Discriminate (FLD).

#### A. Support Vector Machines

SVM is based on distribution n-directional data by n-1 dimensional separation hyperplane. SVM searches for optimal border, based on input data SVM, to separate objects to correct class. Upper mentioned hyperplane is used to do this by same division probabilities as training data. Problem of the training model is to find this optimal border by witch cover and stego characteristic features are divided [4].

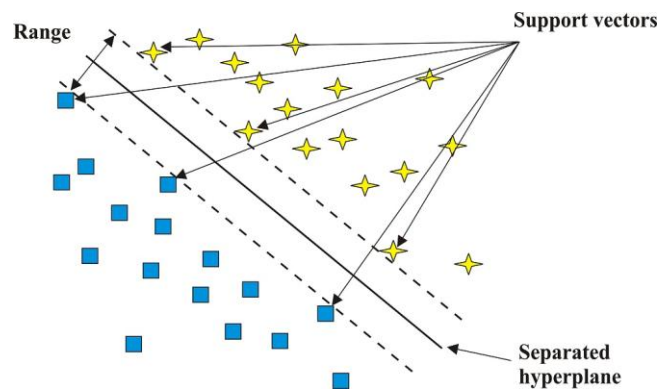


Fig. 3. Linear separated problem classified by SVM. Vectors are classified into two classes divided by separated hyperplane.

Separated hyperplane is defined as:

$$w^T x + b = 0, \quad (1)$$

where  $x$  is input vector,  $w$  is vector of weighting coefficients and  $b$  is an offset.

This case is for linear separated problems. If the problem is not linearly separated, input vector is transformed to space with more dimensions.

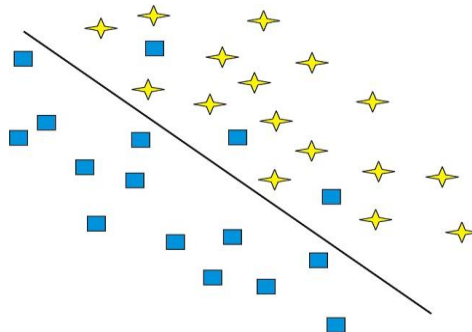


Fig. 4. Linear non-separated problem. In this case input vectors of features are transformed to multidimensional space.

Classifier SVM searches for optimal separated plane in multidimensional space. Separated hyperplane in more-dimensional space is defined:

$$w^T \Phi(x) + b = 0 \quad (2)$$

where  $\Phi(x)$  represents more space dimensions.



### B. Ensemble Classifier

Improved efficiency of individual classifiers is achieved by combination models, consisting of more individual classifiers. Classifier made by this technology is called Ensemble Classifier. Some techniques of combined models are:

- Bootstrap Aggregation
- Boosting

#### 1) Bootstrap Aggregation (bagging)

Ensemble classifier consists of many base learners BL independently trained on a subset dsub of features dimensions dim of input cover and stego images. Each base learner is a simple classifier working with (uniformly) randomly selected subspace of the feature space. In testing phase, each base learner produce final decision whether testing subset of features belongs to cover or stego class. Final decision is made by aggregating of each minority decision [5]. This technique is known as bootstrap aggregation or bagging.

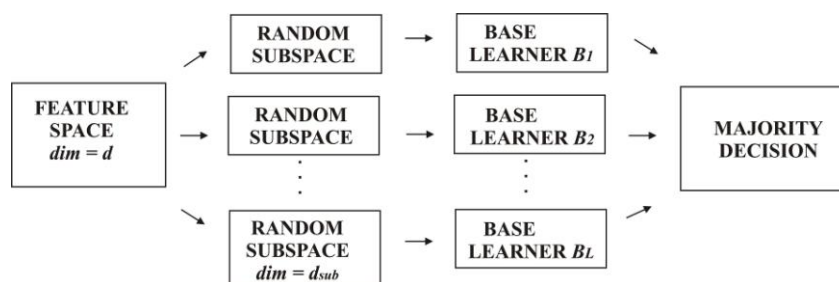


Fig. 5. Scheme of ensemble learning. Random subspaces are obtained from feature space  $d_{im}$  and randomly and uniformly divided into each  $d_{sub}$ .

#### 2) Boosting

Other method of combined learning is boosting. Boosting combine many weak learners (base learners) to make one strong with high accurate. One of the earliest boosting frameworks is AdaBoost. AdaBoost, in compare with bagging, trains individual base learners sequentially and each next base learner with those samples that were more difficult to classify by base learners before. Every base learner focuses on samples that were more difficult to classify by previous base learners. The final decision is achieved by majority decision based on minority decisions of all base learners. This method achieves high accurate and high efficiency of training time [5].

## IV. EXPERIMENTAL RESULTS

In our work we focused on efficiency of machine learning using individual classifier in comparison to ensemble method. From two ensemble methods mentioned in previous section we decided for bagging. As base learners we used FLD (Fisher Linear Discriminate). FLD is relatively simple classifier with fast algorithm appropriate to bagging method that uses weak learners to make one strong. Vector of features extracted consist of 548 features obtained from DCT domain (CC-PEV features).

TABLE I shows the training times of both SVM and Ensemble learners depending on increase number of training samples. To train we used stego images with secret message embedded by steganography algorithm MB2 with payloads: 0.01, 0.02, 0.05, 0.07, 0.1, 0.15, 0.2 and 0.25. Training time of SVM and Ensemble classifiers was diametrically different. With 1,000 samples training time by SVM took 2 minutes while training with same conditions by Ensemble learner took 5 seconds. SVM training with 12,000 samples was 6.5 hours long in comparison to 1 minute by Ensemble. Even 28,000 samples Ensemble trained in 2.25 minutes.

Table 1  
 Training time depending on the  $N^{trn}$  using SVM and ENSEMBLE CLASSIFIER. Steganographic algorithm – MB2.

$N^{trn}$	SVM (trn. time)	Ensemble (trn. time)
1,000	10 m	10 s
2,000	1 h	< 30 s
3,000	2 h 30 m	< 30 s
4,000	5 h	< 30 s
5,000	6 h 30 m	< 1 m
6,000	12 h	< 1 m

In the TABLE II testing results are shown obtained by Ensemble and SVM classifier depend on testing images payloads. Payload is defined in bits per non zero AC DCT coefficients (bpnz) and steganography algorithm used to embed secret information was nsF5. Used model was trained from 4,000 cover and stego images with various resolutions, noises and payloads. In testing phase we used 200 images with various resolutions and noises, not same images as in training phase. While payload was increasing accuracy of detection stego image was improving. With payload 0.25 bpnz was over 60 % images detected and set with payload 0.5 bpnz was already over 90 %. Accuracy of SVM is slightly better, but comparable with Ensemble.

Table 2  
 Accuracy of ensemble detection. Steganographic algorithm – nsF5

<i>payload</i> [bpnz]	SVM [%]	Ensemble [%]
0.01	16	15
0.02	16	16
0.05	16.5	16.5
0.1	21	20
0.25	63	62.5
0.5	94.5	94
0.75	98.5	97
1	100	100

The TABLE III shows similar dependence as TABLE II, but with MB2 steganography algorithm used. Detection accuracy increases with increasing payload (as in previous table). Sufficient decision classifier achieve already in payload equal 0.07 bpnz (0.05 bpnz with Ensemble already). From the table we can judge that detection of MB2 algorithm is slightly better with using Ensemble classifier.

Table 3  
 Accuracy of ensemble detection. Steganographic algorithm – MB2

<i>payload</i> [bpnz]	SVM [%]	Ensemble [%]
0.01	25.5	27.5
0.02	34.5	37.5
0.03	44.5	51
0.05	67	80
0.07	84.5	91
0.1	92.5	95.5
0.15	98.5	99.5
0.25	100	100

## V. CONCLUSION

Modern methods of features based steganalysis are starting to use techniques combining more individual weak learners to make one strong. As was shown in TABLE I, training time is diametrically less than time of individual learner (SVM). The mentioned gives possibility to train very complex models from many input features.

Accuracies of Ensemble and SVM classifiers are comparable. Detection of nsF5 algorithm was slightly better with SVM, while detection of MB2 was better with Ensemble.

## REFERENCES

- [1] M. Filo, "Steganographic Systems and Digital Watermarking for Digital Vector Maps," (In Slovak) diploma thesis, Departments of Informatics, Faculty of Mathematics, Physics and Informatics, Comenius University, Slovak Republic: Bratislava, 2005.
- [2] K. Machová, *Machine learning*. (In Slovak), Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University, Košice, 2002, pp. 9.
- [3] I. J. Cox et al., *Digital Watermarking and Steganography*. 2nd ed., Morgan Kaufmann Publishing, USA: Burlington, 2008, ch.12-13, ISBN: 978-0-12-372585-1.
- [4] D. Majerčák, *Steganalysis of Static Picture based on DCT Transform*. (In Slovak), Faculty of Electrical Engineering and Informatics, Technical University, Slovak Republic: Košice, 2012, ch. 4.
- [5] J. Kodovský, J. Fridrich, and V. Holub, *Ensemble Classifiers for Steganalysis of Digital Media*. IEEE Transactions on Information Forensics and Security, Vol. 7, No. 2, pp. 432-444, April 2012.

# Metódy oceňovania prenosových služieb

<sup>1</sup>Martin PLOTH, <sup>2</sup>Alexander MÉSZÁROS, <sup>3</sup>Ján ZBOJOVSKÝ

Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>martin.ploth@student.tuke.sk, <sup>2</sup>alexander.meszaros@tuke.sk,  
<sup>3</sup>jan.zbojovsky@tuke.sk

**Abstrakt** — Práca popisuje problematiku oceňovania prenosových služieb. Teoreticky je analyzovaná metóda poštovej známky, metóda zmluvnej cesty, metóda MW-km, metóda nevyužitej prenosovej kapacity a metóda distribučných koeficientov. Na modeli 8-zbernicovej sústavy je za pomoci metódy distribučných koeficientov, a to konkrétne GLDF faktora vypočítaná nákladovosť prevádzky jednotlivých vetiev.

**Kľúčové slová** — metóda distribučných koeficientov, metóda MW-km, metóda nevyužitej prenosovej kapacity, metóda poštovej známky, metóda zmluvnej cesty

## I. ÚVOD

Jedným z najzložitejších problémov vyplývajúcich z reštrukturalizácie trhu s elektrickou energiou je spravodlivé prerozdelenie nákladov prenosu medzi každú realizovanú transakciu. Každý realizovaný prenos v prenosovej sústave má dôsledky na každého účastníka v sústave, preto je náročné prerozdeliť náklady na prenos tak, aby kopírovali reálne fyzikálne toky a boli aj jednoduché na vyčíslenie. Existuje veľké množstvo metód, ako k problematike oceňovania pristupovať. Na územi SEPS, a.s., prevádzkovej v SR, sa v súčasnosti používa metóda poštovej známky, ktorá náklady prevádzky prenosovej sústavy prerozdeľuje medzi každú jednu transakciu, len s uvažovaním veľkosti prenášaného výkonu. Avšak táto metóda nerozdeľuje náklady objektívne medzi každý realizovaný prenos. Aby metóda oceňovania prenosu bola spravodlivá, musí rešpektovať fyzikálne toky sústavou. Priradenie činných strát v každej vetve sústavy je zložitá úloha, pretože sa vyžaduje ich prerozdelenie medzi generátory a záťaž, kde nexistuje reálne meranie, alebo matematická metóda, ktorá by prerozdelila náklady rovnakým spôsobom. V tomto článku je opísaných niekoľko metód výpočtu na prerozdelenie nákladov, ako príklad prerozdelenia nákladov je podrobnejšie rozoberaná metóda distribučných faktorov.

## II. METÓDY OCEŇOVANIA PRENOSOVÝCH SLUŽIEB

Každý prístup k problematike oceňovania prenosových služieb musí brať v úvahu rozloženie nákladov na fixné a variabilné. Pri určovaní ceny prenosových služieb je potrebné brať v úvahu taktiež špecifické vlastnosti komodity. Z cenového hľadiska existujú tri základné charakteristiky: úspory z veľkovýroby, funkcie nelineárneho toku elektrickej energie a externalít [1].

Vychádzajúc zo základných cenových charakteristík definujeme prenosové služby dvoma spôsobmi, a to ako Point to point (PTP), alebo ako Branch base (BB).

Na jasné vyčíslenie nákladov na prenos množstva energie sa používa metóda PTP. Nevýhodou metódy však je, že nevyčísľuje dostatočne náklady každej vetvy elektrickej siete a to z dôvodu vplyvu externalít. Cieľom metódy je stanovenie nákladov každej prenosovej služby. Náklady každej vetvy a uzla je možné jasne vyčísliť metódou BB vzhľadom na nelinearitu toku energie, avšak je problematické určiť prenos množstva elektrickej energie. Metóda poskytuje jasný nákladový prehľad v rámci sústavy. Tak pri PTP, ako aj pri BB metóde môžu kalkulované náklady predstaviť celkové náklady, priemerné náklady, alebo hraničné náklady. Podľa typu transakcie rozdeľujeme metódy oceňovania na základe bilaterálneho prenosu BTB (bilateral transmission based) využívanej v cezhraničnom prenose a na základe uzlov NB (node based) aplikovanej obvykle v rámci prenosovej sústavy.

Okrem nákladov na prenášajú elektrickú energiu je potrebné zahnúť aj náklady na rezervovanú kapacitu. Pokiaľ metóda zahŕňa aj náklady na rezervovanú kapacitu v prenášaných MWh, jedná sa o jednotarifnú metódu oceňovania. Naopak pri dvojtarifných metódach reálne využitie prenosovej cesty nemá vplyv na určené náklady rezervovanej kapacity.

#### A. Metóda poštovej známky (Postage stamp method)

Táto metóda vychádza z predpokladu, že prenosová sústava je jedno zariadenie a náklady na jeho prevádzku sa prerozdedia medzi všetky realizované transakcie, bez uváženia rozsiahlosti využitia prenosovej sústavy a fyzikálnych tokov pri každej transakcii, len na základe množstva prenášaného výkonu [3]. Pri tejto metóde je možné jednoducho a jasne odhadnúť náklady na každú transakciu. Nevýhody tejto metódy sa snažia odstrániť jej modifikovaním, napríklad Tracking – based postage – stamp metóda [2], ktorá berie v úvahu aktuálne príspevky generátorov v sústave a zmeny zaťaženia, avšak nezohľadňuje lokálne zaťaženie. Takto modifikovaná metóda poštovej známky odráža spravodlivejšie prerozdelenie nákladov na prenos, ktoré sú nižšie ako pri štandardnej metóde.

#### B. Metóda zmluvnej cesty (Contract path method)

Metóda je postavená na dohodnutej zmluvnej ceste [3] medzi bodom vstupu a výstupu dodávky v prenosovej sieti. Pri wheelingovej transakcii sa dodávateľ a odberateľ dohodnú na prenosovej ceste, ktorá predstavuje súhrn prenosových zariadení, ktoré sa pri oceňovaní budú brať v úvahu. Štandardne sa cesta určuje bez vykonanie štúdie o toku prúdov cez prenosové zariadenia využité pri transakcii. Podiel nákladov je rozdelený medzi transakcie v pomere k využitiu prenosových ciest. Nevýhodou metódy je fakt, že zákazníci môžu dostať nesprávne ekonomické signály a teda výpočet nákladov je potencionálne diskriminačný voči ostatným užívateľom. Prínosom tejto metódy je, že je jednoduchá a stabilné cenové pravidlá sú ľahšie interpretovateľné, a tak metóda lepšie odzrkadľuje využívanie prenosových ciest, ako metóda poštovej známky.

#### C. Metóda MW-kilometer

Metóda nazývaná tiež ako metóda vedenie po vedení [4] (line-by-line) berie v úvahu zmeny prenosových tokov a dĺžku prenosového vedenia. Táto metóda sa často používa pri oceňovaní cezhraničných prenosov. Prenosové poplatky sú stanovené ako funkcia prenášaného objemu energie, dohodnutej cesty a prenosovej vzdialenosti. Pri využití tejto metódy je potrebné realizovať výpočet tokov jednosmerných výkonov. Nevýhodou tejto metódy je fakt, že neberie v úvahu účinník zaťaženia, čiže prenášaný jalový výkon. Nedostatky tejto metódy je možné odstrániť použitím metódy nevyužitej rezervovanej kapacity. MW-kilometer metóda je prvá stratégia stanovenia cien, ktorá reflektuje fixné náklady prenosu vzťahujúce sa na reálne využitie prenosovej cesty.

Pri stanovovaní nákladov sa postupuje nasledovne:

- 1) pre každú transakciu  $t$  sa na základe dodávky výkonu do uzla sa vypočítajú toky výkonu vzťahujúce sa na transakciu v každej vetve siete s aplikovaním jednosmerného modelu prietoku výkonu. Veľkosť toku MW každej vetvy sa vynásobí dĺžkou prenosu a nákladmi €/MW-kilometer a súčin sa vyčísli pre všetky vedenia. Tento výpočet sa vykoná pre každú transakciu.
- 2) príspevok transakcie  $t$  do celových nákladov prenášaného výkonu sa proporcionálne prerozdeli pomerne k veľkosti toku podieľajúcej sa na transakcii  $t$  a súčtu absolutných tokov spôsobených každou transakciou, ako je to vyjadrené vzťahom

$$TC_t = TC \cdot \frac{\sum_{k \in K} c_k L_k MW_{t,k}}{\sum_{t \in T} \sum_{k \in K} c_k L_k MW_{t,k}} \quad (1)$$

kde

- $TC_t$  sú náklady pridelené na transakciu  $t$ ,
- $TC$  - celkové náklady všetkých vedení v €,
- $L_k$  - dĺžka vedenia  $k$  v kilometroch,
- $c_k$  - náklady na MW na jednotku dĺžky vedenia  $k$ ,
- $MW_{t,k}$  - tok vo vedení  $k$ , patriaci k transakcii  $t$ ,
- $T$  - séria transakcií,

$K$  - séria vedení.

#### D. Metóda nevyužitej prenosovej kapacity

Metóda je založená na princípe, ktorý berie v úvahu to, že používateľ prenosovej cesty má platiť rovnako ako za využitú, tak aj za nevyužitú prenosovú kapacitu [5], [6]. Tento prístup zabezpečí prevádzkovateľovi prenosovej sústavy celkovú návratnosť prenosu nezávisle od využitia prenosovej kapacity. Problémom metódy je fakt, že z dlhodobého hľadiska neprináša pozitívne stimuly rozvoja a efektívneho využívania prenosovej sústavy. Ďalšou nevýhodou je fakt, že pokiaľ je linka predimenzovaná z dôvodu zabezpečenia spoľahlivosti prenosovej sústavy a za štandardnej prevádzky málo vytážená, je používateľ nútený platiť neprímerané poplatky. Tento nedostatok je možné odstrániť parciálnym prerozdelením nákladov na takejto linke. Avšak metóda nezohľadňuje vplyv prebytku kapacity, ktorý má dopad na spoľahlivosť prenosovej sústavy.

Metóda je určená ako rozdiel medzi výkonovými kapacitami zariadení a ich reálnym zaťažením. Tento rozdiel sa nazýva nevyužitá prenosová kapacita a je vyjadrená vzťahom

$$TC_t = \sum_{k \in K} C_k \frac{|F_{t,k}|}{\sum_{t \in T} |F_{t,k}|} \quad (2)$$

kde

$TC_t$  sú náklady pridelené na transakciu  $t$ ,  
 $C_k$  - vložené náklady zariadenia  $k$ ,  
 $|F_{t,k}|$  - výkonové zaťaženie zariadenia spôsobené transakciou  $t$ ,  
 $T$  - séria transakcií,  
 $K$  - séria vedení.

#### E. Metóda distribučných koeficientov

Alternatívou k uvedeným metódam je metóda distribučných koeficientov [7]. Vychádza z lineárnych metód riešenia elektrických sietí. Pôvodne boli tieto koeficienty využívané v analýzach bezpečnosti a rizika. Avšak je možné ich využiť tiež ako mechanizmus pridelovania výšky poplatku za prenos. Distribučné koeficienty môžu byť vyjadrené nasledovne:

1. Generation Shift Distribution Factors (GSDFs koeficienty alebo tzv. A koeficienty) stanovujú zmeny výkonových tokov vo vedeniach v dôsledku zmien v generovanom výkone do sústavy. Využívajú sa pri stanovení maximálnych tokov prenosu pre danú výrobu a zaťaženie, je definovaný nasledovne:

$$\Delta F_{l-k} = A_{l-k,i} \Delta G_i \quad (3)$$

$$\Delta G_r = -\Delta G_i \quad (4)$$

kde

$\Delta F_{l-k}$  je zmena činného výkonu medzi uzlami  $l$  a  $k$ ,  
 $A_{l-k,i}$  - A koeficient (GSDF) vedenia spájajúceho uzly  $l$  a  $k$ , odpovedajúci zmene výroby v uzle  $i$ ,  
 $\Delta G_i$  - zmena vo výrobe v uzle  $i$ , okrem bilančného uzla,  
 $\Delta G_r$  - zmena vo výrobe v bilančnom (referenčnom) uzle  $r$ .

2. Generalized Generation Shift Distribution Factors (GGDFs koeficienty alebo tzv. D koeficienty) určujú pôsobenie každého generátora v sústave na toky činného výkonu. Stanovujú celkové využitie zariadení prenosovej sústavy. Koeficienty sú založené na modeli jednosmerného prúdu a sú definované nasledujúco

$$F_{l-k} = \sum_{i=1}^N D_{l-k,i} G_i \quad (5)$$

$$D_{l-k,i} = D_{l-k,r} + A_{l-k,i} \quad (6)$$

$$D_{l-k,r} = \left\{ F_{l-k}^0 - \sum_{\substack{i=1 \\ i \neq r}}^N A_{l-k,i} G_i \right\} / \sum_{i=1}^N G_i \quad (7)$$

kde

- $F_{l-k}$  je celkový činný tok medzi uzlami  $l$  a  $k$ ,
- $F_{l-k}^0$  - tok výkonu medzi uzlami  $l$  a  $k$  z predchádzajúcej iterácie,
- $D_{l-k,i}$  - D koeficient (GGDF) vedenia medzi uzlami  $l$  a  $k$ , odpovedajúci generátoru v uzle  $i$ ,
- $D_{l-k,r}$  - koeficient GGDF vedenia medzi uzlami  $l$  a  $k$ , v dôsledku generovania výkonu v referenčnom uzle  $r$ ,
- $G_i$  - celkové generovanie výkonu v uzle  $i$ .

3. Generalized Load Distribution Factors (GLDFs alebo C koeficienty) stanovujú podiel každého zaťaženia na tok vedenia. Slúžia na vyčíslenie poplatkov od podriadenej prenosovej siete na zaťaženie v prevádzke rozvodnej siete. Sú definované takto:

$$F_{l-k} = \sum_{j=1}^N C_{l-k,j} L_j \quad (8)$$

$$C_{l-k,j} = C_{l-k,r} - A_{l-k,j} \quad (9)$$

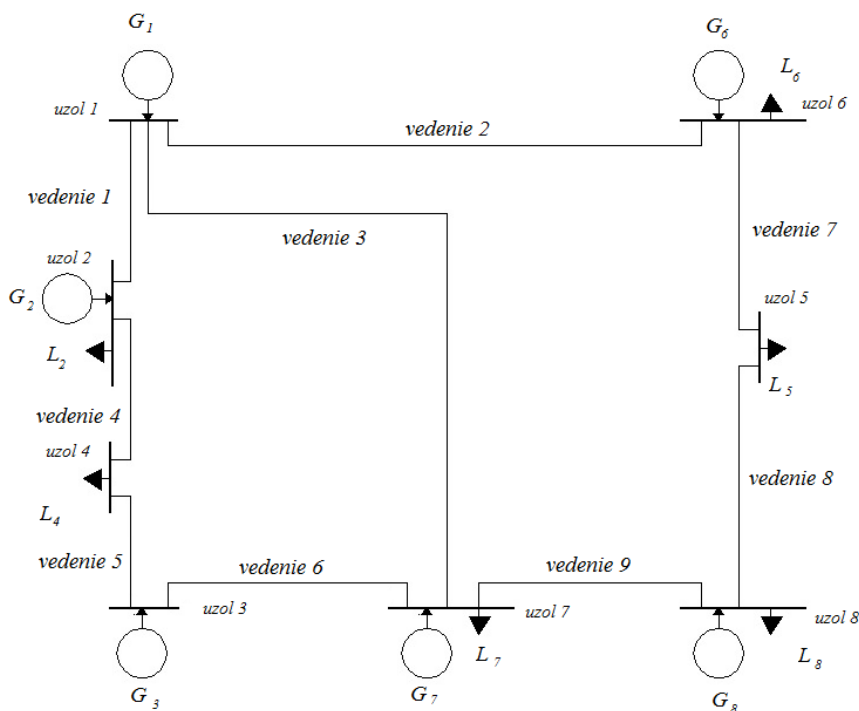
$$C_{l-k,r} = \left\{ F_{l-k}^0 - \sum_{\substack{j=1 \\ j \neq r}}^N A_{l-k,j} L_j \right\} / \sum_{j=1}^N L_j \quad (10)$$

kde

- $F_{l-k}$  - celkový tok činného výkonu medzi uzlami  $l$  a  $k$ ,
- $F_{l-k}^0$  - tok výkonu na vedení medzi uzlami  $l$  a  $k$  z predchádzajúcej iterácie,
- $C_{l-k,i}$  - C koeficient (GLDF) vedenia medzi uzlami  $l$  a  $k$ , odpovedajúci generátoru v uzle  $j$ ,
- $C_{l-k,r}$  - koeficient GLDF pre vedenie medzi uzlami  $l$  a  $k$ , v dôsledku generovania výkonu v referenčnom uzle  $r$ ,
- $L_i$  - celkový odber v uzle  $i$ .

### III. OCEŇOVANIE METÓDOU DISTRIBUČNÝCH KOEFICIENTOV

V tomto odseku na základe vzorového príkladu, ktorý predstavuje 8-zbernicový systém, sú vypočítané prerozdelenia nákladov za pomoci využitia metódy distribučných koeficientov, a to konkrétne GLDF faktora, tzv. C koeficientov [8]. Výpočet jednotlivých tokov bol realizovaný za pomoci simulačného softvéru GLF.



Obr. 1. 8-zbernicový systém modelového prípadu s parametrami jednotlivých vedení

Tabuľka 1  
 Parametre vedení

Vedenie č.	od	do	R [p. j.]	X [p. j.]	B [p. j.]	L [km]
1	1	2	0,00415	0,025	0,04	30
2	1	6	0,0097	0,0058	0,095	70
3	1	7	0,0166	0,101	0,161	120
4	2	4	0,00415	0,025	0,04	30
5	4	3	0,00623	0,0376	0,06	45
6	3	7	0,00526	0,032	0,052	38
7	6	5	0,00415	0,025	0,04	30
8	5	8	0,00279	0,017	0,027	20
9	8	7	0,00554	0,0334	0,0538	40

Tabuľka 2  
 Charakteristické hodnoty výroby a zaťaženi v jednotlivých uzloch sústavy

Zbernica	U [kV]	P <sub>C</sub> [MW]	Q <sub>C</sub> [MVar]	P <sub>G</sub> [MW]	Q <sub>G</sub> [MVar]
1	231	0	0	490,89	285,76
2	220	300	35	340,49	-122,84
3	220	0	0	350	13,37
4	217,523	300	25	0	0
5	217,399	208	30	0	0
6	220	230	60	291,8	68,64
7	215,44	350	38	0	0
8	214,939	170	20	15,17	-6,16

Tabuľka 3  
 Toky výkonov pre jednosmerný výpočet a číselné hodnoty C faktorov

Vedenie č.	od	do	P <sub>ij</sub> [MW]	Q <sub>ij</sub> [MVar]	C <sub>ik,7</sub>	C <sub>ik,4</sub>	C <sub>ik,5</sub>	C <sub>ik,8</sub>
1	1	2	153,99	184,1	-0,141	0,280709	-0,21762	-0,196
2	1	6	211,305	60,82	-0,1697	-0,3868	0,00189	-0,04549
3	1	7	125,53	40,9	0,004434	-0,2	-0,09039	-0,06445
4	2	4	192,289	17,089	-0,1959	0,22538	-0,2729	-0,252
5	4	3	-109,201	-13,3	-0,25649	0,7459	-0,3043	-0,31991
6	3	7	239,99	1,4	-0,03021	-0,106	-0,25429	-0,2269
7	6	5	233,68	48,5	-0,2316	-0,31733	-0,0705	-0,1213
8	5	8	23,4	8,145	0,07249	-0,0132	-0,546	0,24009
9	8	7	-173,51	-3,701	-0,4272	-0,3131	0,09657	0,1806

Náklady na prenos boli stanovené na rovnakej hodnote, ako cena za prenos kapacity stanovená rozhodnutím ÚRSO pre SEPS a.s. a to pre zjednodušený výpočet zaokrúhlené na hodnotou 1,74 €/MWh.

Tabuľka 4  
 Celkové náklady na prenos elektriny

Vedenie č.	Merné náklady na prenos [€/MWh]	P <sub>ij</sub> [MW]	Celkové náklady na prenos [€]
1	51	113,29	5777,79
2	121,8	192,03	23389,254
3	208,8	113,497	23698,1736
4	51	188,98	9637,98
5	78,3	111,02	8692,866
6	66,12	238,991	15802,0849
7	51	228,64	11660,64
8	34,8	20,49	713,052
9	69,6	175,41	12208,536
Celkovo			111580,3765

#### IV. ZÁVER

Bilančná odchýlka musí byť ocenená tak, aby pokrývala vzniknuté náklady pri prenose elektriny. Cena za bilančnú odchýlku je cenou regulovanou s ÚRSO. Zúčtovateľ môže byť prenosová sústava, alebo nezávislá spoločnosť, v závislosti od platnej legislatívy.

V tejto práci sú uvedené výsledky z riešenia modelového príkladu s 8-zbernicovou sústavou podľa obr.1. Na výpočet nákladov na prenos v každej vetve bola použitá rozšírená metóda distribučných koeficientov. Rozloženie nákladov založené na tejto hypotéze dostatočne presne odzrkadľuje reálne náklady prenosu elektriny na jednotlivých sieťových prvkoch. Možno predpokladať, že zmena činného výkonu generovaného nezaťaženými zbernicami nemení

číselnú hodnotu distribučných faktorov. Pri tejto metóde výpočtu je potrebné brať v úvahu činné výkonové straty.

#### POĎAKOVANIE

Túto prácu podporila Vedecká grantová agentúra Ministerstva školstva SR a Slovenskej akadémie vied projektom VEGA 1/0388/13.

#### LITERATÚRA

- [1] Zhao-xia Jing; Xian-zhong Duan, "The cost and pricing method of uniform transmission services," *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, vol.3, no., pp.1787-1791 vol.3, 2002
- [2] Radzi, N. H.; Bansal, R.C.; Dong, Z.Y.; Hassan, M.Y., "A modified postage-stamp coverage method for local load case of transmission service charges," *Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011 4th International Conference on*, vol., no., pp. 378-385, 6-9 July 2011
- [3] Kranthi Kiran, I; Jaya Laxmi, A.; Power flow based contract path method for transmission pricing , "International Journal of Soft Computing and Engineering" (IJSCE), volume-3, issue-6, pp.61-65, 2014, ISSN 2231-2307
- [4] Shahidepour, Mohammad; Yamin, Hatim; Li, Zuyi, "Market Operations in Electric Power Systems ", Library of congress Cataloging in Publication Data, 2002, ISBN 0-47L-44337-9
- [5] Arabali, A.; Hosseini, S.H.; Moeini-Aghaie, M., "Pricing of transmission services: An efficient analysis based on fixed and variable imposed costs," *Environment and Electrical Engineering (EEEIC), 2012 11th International Conference on* , vol., no., pp. 407-412, 18-25 May 2012
- [6] Chunyang Zhao; Galloway, S.; Lo, K.L., "A new method for charging transmission services," *Universities Power Engineering Conference (UPEC), 2009 Proceedings of the 44th International*, vol., no., pp.1-5, 1-4 Sept. 2009
- [7] Qian Zhang; Jihui Yu, "Research on transmission services pricing based on allocation of fixed transmission cost," *Services Systems and Services Management, 2005. Proceedings of ICSSSM '05. 2005 International Conference on* , vol.2, no., pp.1369-1374 Vol. 2, 13-15 June 2005
- [8] Kilyeni, S.; Pop, O.; Slavici, T.; Craciun, C.; Andea, P.; Mnerie, D., "Transmission cost allocation using the distribution factors method", MELECON 2010 -2010 15th IEEE Mediterranean Electrotechnical Conference, pp 1093-1098, 26-28 April 2010, ISBN 978-1-4244-5793-9



# Analýza veľkých dát použitím technológie IBM InfoSphere BigInsights

Adriána ŠANTOVÁ, Peter MICHALIK, Iveta ZOLOTOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

adriana.santova@student.tuke.sk, peter.michalik.2@tuke.sk, iveta.zolotova@tuke.sk

**Abstract** — Veľké dáta predstavujú veľké množstvo dát, ktoré sú náročné na spracovávanie podľa tradičných databázových systémov. Tento článok sa zameriava na analýzu dát, ktorá môže firme pomôcť získať konkurenčnú výhodu alebo zlepšiť rozhodovacie procesy vo firme. V článku sú popísané základné informácie týkajúce sa vývoja analýzy dát, základné koncepty Business Intelligence. Hlavné zameranie článku sa týka oblasti analýzy dát novovznikajúceho trendu Big Data. Popisuje jej základné charakteristiky a technológie, ktoré používa pri svojej práci. Záver tejto práce obsahuje popis softvéru IBM InfoSphere BigInsights.

**Keywords** — Analýza dát, Business Intelligence, Big Data, IBM Infosphere BigInsights

## I. ÚVOD

Na trhu existuje množstvo organizácií, ktoré sa neustále snažia zvyšovať predajnosť svojich výrobkov a služieb. Tieto podniky medzi sebou súťažia, každý z nich chce byť na trhu najlepším a to sa snažia dosiahnuť rozličnými spôsobmi.

Jedným zo spôsobov je získavanie dát od zákazníkov a následne ich analýza, prostredníctvom ktorej firmy získavajú informácie o potrebách zákazníkov a tým zvyšujú svoju podnikateľskú činnosť. Ďalšou možnosťou zlepšenia postavenia organizácie na trhu je zdokonalenie rozhodovacích procesov.

Keďže množstvá dát pribúdajúcimi rokmi narastali, tak vznikali nové spôsoby analýzy dát. Analýza dát sa postupom času zdokonaľovala od obyčajných popisných štatistík až po súčasnú podnikovú inteligenciu. Vývoj však nezastal a neustále sa zlepšujú existujúce spôsoby analýzy dát. Najnovšie princípy analýzy dát v dôsledku ich vývoja predstavujú Big Data, t.j. veľké dáta.

## II. BUSINESS INTELLIGENCE

Koncept Business Intelligence nie je úplne nový. Mnohé organizácie sa pokúšali vsunúť svoje systémy do služieb na produkciu lepších strategických pohľadov.

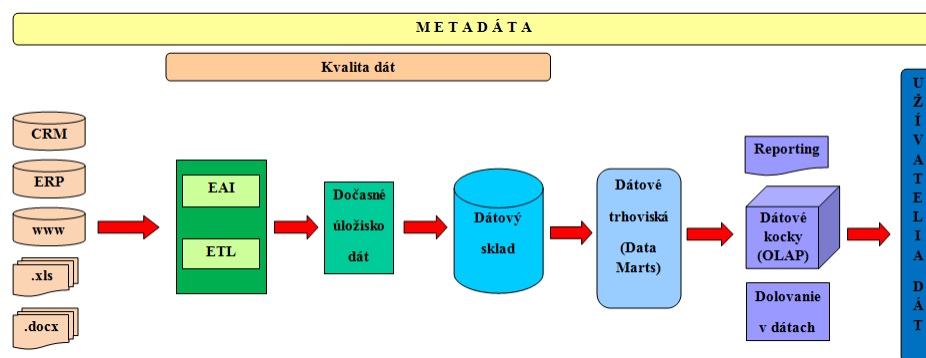
V posledných pár desaťročiach vzniká monumentálny prechod týkajúci sa vytvárania, zhromažďovania a využívania informácií. Veľké pokroky boli v oblasti manažmentu znalostí, ale taktiež sa nikdy nezabudlo na obchodnú hodnotu skrytú v hromadách bitov a bajtov. Tieto údaje sú v súčasnosti zdieľané nielen prostredníctvom štruktúrovaných dátových systémov, ale sa zachytávajú, zdieľajú, manažujú, ukladajú a distribuujú prostredníctvom rozličných súborových formátov reprezentovaných štruktúrovanými, ale aj neštruktúrovanými dátami. Práve vďaka tejto obchodnej hodnote vznikol pojem Business Intelligence [1].

Business Intelligence (BI) predstavuje súbor teórií, metód, procesov, architektúr a technológií, ktoré menia dáta na znalosti určené na obchodné účely. Tieto dáta môžu byť získavané z rôznych zdrojov, štruktúrované rôznymi spôsobmi a môžu sa získavať v rozličných množstvách a kvalitách [2].

Je to nástroj, ktorý firmy využívajú na riešenie problémov alebo umožňuje dosiahnuť niečo čo je v súčasnej situácii daného podniku nedosiahnuteľné. Medzi bežné funkcie technológie BI zaraďujeme *reporting*, *online analytické spracovávanie*, *dolovanie v dátach*, *obchodné riadenie výkonnosti*, *analytika* a *benchmarking*, *dátové sklady* [3].

Architektúra podnikovej inteligencie sa skladá zo štyroch hlavných vrstiev [3]:

- prezentačná vrstva – na tejto vrstve vieme prehľadávať všeobecné informácie, ktoré sú analyzované pomocou BI systému. Zaraďujeme sem údaje z CRM, ERP systémov, dokumenty a mnohé iné zdroje.
- vrstva extrakcie, transformácie a čistenia dát – je to vrstva, kde dochádza k zberu a prenosu dát z produkčných systémov do vrstvy na ukladanie dát BI riešení. Tu patria transformačné nástroje ETL a EAI.
- vrstva ukladania dát – na tejto vrstve sa vykonávajú procesy ako ukladanie, aktualizácia a správa dát ohľadom riešení BI. Pričom zaraďujeme sem nástroje ako dočasné úložisko dát, dátové sklady a dátové trhoviská.
- vrstva analýzy dát – táto vrstva obsahuje analýzy dát a to na základe reportingu, dátových kociek a dolovania v dátach.



Obr. 1 Architektúra Business Intelligence

### III. BIG DATA

Rast množstva dát nezastal, analýzy dát sa taktiež zdokonaľujú, došlo k vzniku pojmu Big Data. Dôvodom vzniku pojmu Big Data je, že dáta sú príliš veľké alebo sa pohybujú príliš rýchlo. Ak chceme získať nejaké dáta, musíme použiť alternatívne spôsoby ich spracovania [4].

Najväčšie výzvy a výhody pre podniky rôznych odvetví predstavovali riadenie a analýza dát bez ohľadu na ich veľkosť. Podniky hľadali rôzne pragmatické prístupy na archiváciu informácií o svojich produktoch, službách a zákazníkoch. Pri menšom počte zákazníkov bolo veľmi jednoduché tieto informácie uchovávať. Časom to už bolo náročnejšie, keďže sa podniky a trh výrobkov a služieb rozrástol a tento vývoj nezastal. S uchovávaním dát nemali problémy len v podnikoch, ale aj vo vede, výskume [5].

Dáta poznáme rôznych typov a to *štruktúrované dáta*, ktoré sú uložené v tradičných relačných databázach. Na druhej strane sú *neštruktúrované dáta*, ako napríklad obrázky, videá, dokumenty, záznamy zákazníckych služieb. Najväčšou výzvou pre podniky v súčasnosti predstavuje možnosť ako dať zmysel priesečníku všetkých typov dát. To predstavuje apel a príležitosť pre BD [5].

#### A. Charakteristika Big Data

Veľké dáta majú tri základné charakteristiky, ktoré sa taktiež označujú ako „3V“. V prípade že sa vzťahujú na kľúčové atribúty veľkých dát, je nutné zaviesť aj štvrtú charakteristiku, tzv. „4V“:

**Objem (volume)** - považuje sa za vlastnosť, ktorá najviac vystihuje podstatu veľkých dát. Poukazuje na obrovské množstvá dát, ktoré podniky využívajú pre lepšie rozhodovanie v rámci celého podniku. Množstvo dát sa však zväčšuje nepredvídaným tempom, ktoré je rozličné v jednotlivých odvetviach. Viaceré podniky už teraz zápasia s neustále sa zväčšujúcou databázou [6].

**Rôznorodosť (variety)** – predstavuje rôzne typy dát a dátových zdrojov, ktoré sa týkajú rozhodovacích procesov. Rozmanitosť sa nachádza v riadení komplexity rôznych typov dát, medzi ktoré zaraďujeme *štruktúrované, pološtruktúrované*, ale aj *neštruktúrované* údaje. Organizácie potrebujú začleniť a analyzovať dáta z komplexných radov tradičných aj netradičných informácií v rámci podniku aj mimo neho. NoSQL poskytuje dostatok štruktúry na organizovanie dát a nevyžaduje pri tom výber schémy pred uložením dát [7]. V súčasnosti sa dáta vytvárajú v nespočetných formách, ako napríklad audio, video, text, uzavreté súbory, zvuk, dokumenty, obrázky a mnohé ďalšie. Keďže tieto dáta sú veľmi rozmanité je náročné jednotlivé dáta zaradiť do skupín, do ktorých patria [6].

**Rýchlosť** (*velocity*) – popisuje rýchlosť akou sú dáta vytvárané, spracovávané a analyzované stále sa zvyšujúcou rýchlosťou. Zvyšuje sa rýchlosť, ktorá je v danom čase charakteristická pre tvorbu dát aspoň tak, ako si to vyžaduje začlenenie toku dát do podnikového procesu a rozhodovacieho procesu [7].

**Pravdivosť** (*veracity*) – sa týka úrovne spoľahlivosti, s ktorou sú spojené určité typy dát. Úsilie mať kvalitné dáta je veľmi dôležité. Toto úsilie predstavuje veľkú výzvu, ale ani tie najlepšie čistiace metódy nemôžu odstrániť nepredvídateľnosť niektorých dát ako napríklad počasie, ekonomika alebo rozhodovanie zákazníka o budúcej kúpe tovaru alebo služby [6].

### B. *Technológie Big Data*

**Hadoop** predstavuje najpragmatickejší spôsob, ktorý umožňuje spravovať obrovské objemy dát veľmi ľahko. Umožňuje rozdeliť veľké problémy na menšie prvky, tak aby sa mohla vykonať analýza rýchlo a efektívne. Rozdrobením problému veľkých dát na menšie prvky, ktoré by mohli byť spracovávané paralelne, sa môžu spracovávať a preskupovať na malé prvky do súčasných výsledkov. Táto technológia je navrhnutá tak, aby sa jednotlivé operácie vykonávali paralelne cez počítačové uzly na urýchlenie výpočtov a zníženie čakacej doby. Okrem toho má ďalšiu funkciu, a to že je schopná detekovať zmeny, vrátane zlyhania, prispôbiť sa týmto zmenám a pokračovať v činnosti bez prerušenia [5]. Má dva primárne komponenty *Hadoop Distributed File System* (HDFS) a *MapReduce*. Využíva aj ďalšie technológie ako je *Pig*, *Hive*, *HBase*, *Zookeeper*, *Oozie* a iné.

**NoSQL Databázy** sú nástroje, ktoré sú navrhnuté na obchodovanie, ľahké používanie tradičných databáz, pre flexibilitu a výkon, čo predstavuje nutnosť pre vývojárov pri riešení nových problémov, s ktorými sa stretávajú. NoSQL poskytuje mechanizmus pre ukládanie a vyhľadávanie dát, ktoré boli namodelované v iných ako tabuľkových vzťahoch, ktoré sú využívané v relačných databázach. Medzi najznámejšie NoSQL databázy zaradujeme *MongoDB*, *HBase*, *Riak*, *Cassandra*, *CouchDB* a iné [8].

**Cloud computing** je spôsob poskytovania súboru zdieľaných výpočtových zdrojov, ktoré v sebe zahŕňajú aplikácie, výpočty, skladovanie, siete, vývoj, nasadzovanie platformy a taktiež aj obchodné procesy. Môže byť dodávaná ako služba týkajúca sa výpočtového výkonu pre počítačovú infraštruktúru a aplikácie, podnikové procesy založené na dátach. Základom je, že to všetko musí byť vykonávané spoločnými štandardizovanými postupmi a automatizáciou [5]. Cloud computing popisuje rôzne výpočtové spôsoby, ktoré vyžadujú veľké množstvo počítačových spojení, ako napríklad komunikačné spojenie v reálnom čase cez internet. Cloud computing predstavuje viac ako len službu, ktorá je vo funkcii sústreďovania vzdialených dát. Je to súbor metód, ktoré umožňujú organizáciám rýchlo a efektívne pridávať alebo uberať zdroje v reálnom čase. Firmy túto technológiu považujú za základ prostredníctvom, ktorého uskutočňujú svoje obchody [9]. Táto technológia môže mať formy Public clouds, Private clouds, Hybrid clouds.

**Distribúované výpočty** nie sú len distribuované výpočty modelu, pretože zdroje na počítači môžu byť distribuované v mnohých ohľadoch. Môžu distribuovať napríklad sadu programov na rovnakom serveri a využívať služby určené na odosielanie správ, čo im umožní komunikovať a odovzdávať informácie. Umožňuje tiež aby mnohé systémy alebo servery, z ktorých čerpá svoje vedomosti mohli spolupracovať a tým spoločne vyriešiť problém. Dostupné hardvéry a technológie pre firmy, umožňujú vykonať komplexnú analýzu dát na základe automatizácie procesov ako je napríklad vyvažovanie, optimalizácia cez obrovské uzlové bloky. Súčasťou tejto technológie sú vstavané pravidlá, ktoré rozumejú pracovným záťažiam, ktoré sú nevyhnutné na určitých výkonnostných úrovniach. Táto technológia spracuje všetky uzly, akoby bola jedným veľkým bazénom výpočtov, úložných a sieťových aktív. Presúva procesy do jedného uzla bez prerušenia v prípade, ak nejaký uzol zlyhá pri používaní technológie vizualizácie [5].

## IV. IBM INFOSPHERE BIGINSIGHTS

### A. *Nástroj InfoSphere BigInsights*

InfoSphere BigInsights je softvér od firmy IBM, ktorý slúži ako základ pre objavovanie, analyzovanie a vizualizáciu dát, ktoré sú z rôznych zdrojov a sú nespárované. Tento softvér pomáha spracovať a analyzovať objem, rôznorodosť a rýchlosť dát, ktoré predstavujú každodenný vstup do organizácie. Softvér pomáha organizáciám porozumieť a analyzovať masívne objemy neštruktúrovaných informácií tak jednoducho, akoby pracovali len s malým objemom informácií. Základ softvéru je flexibilný a je založený na voľne dostupnom rámci Apache Hadoop, ktorý beží paralelne na dostupnom hardvéri s nízkymi nákladmi. Umožňuje použiť škálovateľnú platformu na analýzu stoviek TB, PB a oveľa väčšie množstvo

nespracovaných dát z rôznych zdrojov. Tak ako rastie množstvo informácií je možné pridávať viac hardvérov na podporu prílivu dát. InfoSphere BigInsights taktiež pomáha vývojárom aplikácií, dátovým vedcom a administrátorom vo svojej organizácii rýchlym tempom vytvárať a nasadzovať svoje vlastné analýzy a zachytávať rôzne pohľady dát. Tieto dáta sú často integrované v existujúcich databázach, dátových skladoch a infraštruktúre podnikovej inteligencie. Užívatelia môžu využívaním tohto softvéru získavať nové pohľady z dát a tým zvyšovať svoje obchodné znalosti [10].

### B. Architektúra InfoSphere BigInsights

InfoSphere BigInsights obsahuje rôzne technológie a funkcie. Kombináciou daných technológií a funkcií pomáhajú a uľahčujú organizácii správu dát od ich vstupu do podniku až po ich analýzu a výsledné riešenia. Daný nástroj je rozšírený o Hadoop open source, ktorý obsahuje rámec pre podnikovú bezpečnosť, kontrolu a integráciu existujúcich dát do dátového skladu, nástroje uľahčujúce prácu a zvyšujúcu produktivitu vývojárov a mnohé iné. Apache Hadoop pomáha organizácii riadiť a analyzovať dáta, ktoré predtým nebolo možné. Tento open source nástroj spolu so súvisiacimi technológiami predstavuje kľúčový prvok nástroja InfoSphere BigInsights.

### C. Analýza dát v nástroji InfoSphere BigInsights

Dáta boli preskúmané a následne analyzované a vizualizované pomocou nástroja InfoSphere BigInsights. Na analýzu a následnú vizualizáciu dát bol použitý tabuľkový nástroj InfoSphere BigSheets. Analyzované boli dáta zhromaždené z Twitteru. Predmetom skúmania boli textové príspevky tzv. tweety na danej sociálnej sieti.



Obr. 2 Výsledná vizualizácia dát

Výsledkom tohto prieskumu dát je stĺpcový graf, ktorý znázorňuje 15-tich užívateľov, ktorí v najvyššej miere posielali tweety na sociálnej sieti Twitter týkajúce sa pojmu Hadoop. Na x-ovej osi sa nachádzajú používatelia, používateľské mená tých, ktorí posielali tweety. Množstvo odoslaných tweetov je zobrazené na y-ovej osi.

## V. ZÁVER

Keďže množstvo dát s pribúdajúcim časom neustále narastá, tak je potrebné neustále zdokonaľovať existujúce systémy na analýzu dát. Dôvodom je získanie konkurenčnej výhody v dnešnom digitalizovanom svete. Tento článok popisuje základné poznatky potrebné na pochopenie dôvodu vzniku a fungovania existujúceho a nového trendu v oblasti dátovej analýzy tzv. Big Data.

## VI. POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja (60%). Podporujeme výskumné aktivity na Slovensku/Projekt je

spolufinancovaný zo zdrojov EÚ. Príspevok bol tiež podporený grantom z KEGA - 021TUKE-4/2012 (40%).

#### LITERATÚRA

- [1] David LOSHIN, Business Intelligence – The savvy manager’s guide, 2nd edition, Morgan Kaufman, 225 Wyman Street, Waltham, MA 02451, USA, ISBN:978-0-12-385889-4.
- [2] Rajiv SABHERWAL, Irma BECCERA-FERNANDEZ, Business Intelligence – Practices, Technologies and Management, ISBN: 978-0-470-46170-9
- [3] Swain SCHEPS, Business Intelligence for Dummies, ISBN: 978-0-470-12723-0
- [4] Robert D. SCHNEIDER, Hadoop for DUMMIES, Special edition – John Wiley and Sons Canada, Ltd, ISBN: 978-1-118-25051-8
- [5] Judith HURWITZ, Alan NUGENT, Dr. Fern HALPER, Marcia KAUFMAN, Big Data for DUMMIES A Willey Brand, John Wiley and Sons, Inc 11 river Street Hoboken,NJ 07030-5774, ISBN: 978-1-118-50422-2
- [6] Analytics: The real-world use of big data – How innovative enterprises extract value from uncertain data, IBM institute for Business Value. Dostupné na internete: [www.ibm.com/iibv](http://www.ibm.com/iibv)
- [7] Paul C. ZIKOPOULOS, Dirk de ROOS, Krishnan PARASURAMAN, Thomas DEUTSCH, David CORRIGAN, James GILES, Harness the Power of Big Data, The IBM Big Data Platform,The McGraw-Hill Companies United States Copyright ISBN: 978-0-07180818-7
- [8] Pete WARDEN, Big Data Glossary, O’Reilly Meia Inc, 1005 Gravenstein Highway North, Sebastopol, CA 05472, ISBN: 978-1-449-31459-0
- [9] Judith HURWITZ, Robin BLOOR, Marcia KAUFMAN, Dr. Fern HALPER, Cloud Computing for Dummies, ISBN: 978-470-4470-8
- [10] IBM InfoSphere BigInsights Information Center, Dostupné na internete: <http://pic.dhe.ibm.com/infocenter/bigins/v2r0/index.jsp>

# Procesná analýza výrobnéj firmy s využitím vybraných technológií

<sup>1</sup>Gabriela KONKOLOVÁ, <sup>2</sup>Peter MICHALIK, <sup>3</sup>Iveta ZOLOTOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

<sup>1</sup>gabriela.konkolova@student.tuke.sk, <sup>2</sup>peter.michalik.2@tuke.sk,  
<sup>3</sup>iveta.zolotova@tuke.sk

**Abstrakt** — Článok sa zaoberá problematikou procesnej analýzy a riadenia podnikových procesov, pričom poukazuje na potrebu a vhodnosť ich využitia aj pre malé podniky. Na demonštráciu tejto problematiky je uskutočnená procesná analýza vo výrobnom podniku, ktorá je zameraná na zlepšovanie procesov a taktiež je vyhodnotená výkonnosť identifikovaných podnikových procesov.

**Kľúčové slová** — zlepšovanie procesov, mapovanie procesov, modelovanie, simulácia, Balanced Scorecard

## I. ÚVOD

Súčasný problém malých podnikov je udržať si svoje postavenie na silne konkurenčnom trhu a neustále zlepšovať svoje podnikanie aj napriek tomu, že tento druh podnikov nedisponuje takým veľkým kapitálom, aby mohol investovať do moderných zariadení. Vhodné je preto sústrediť sa na efektívne využívanie vlastných zdrojov [1].

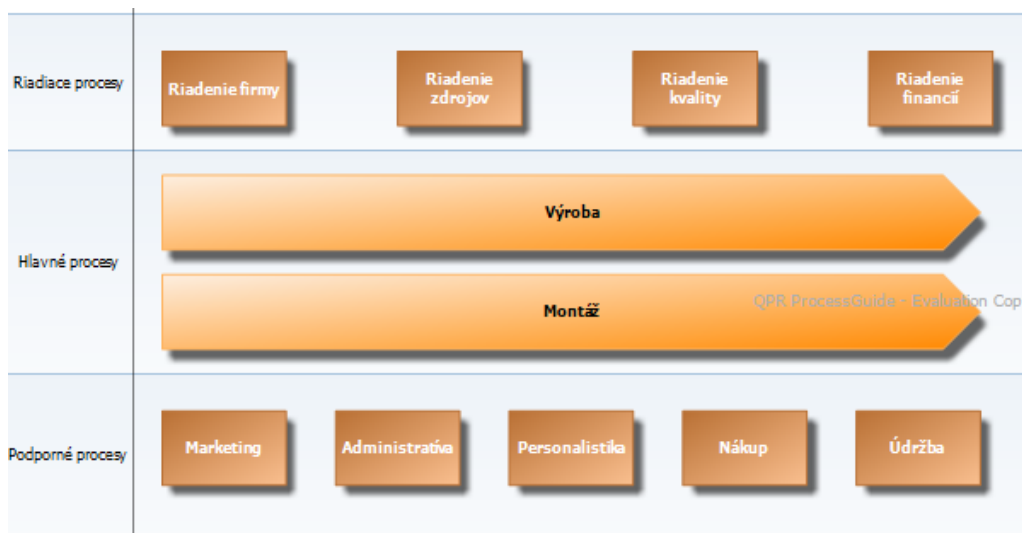
Zmena riadenia z funkčného na procesné je neodmysliteľnou súčasťou zlepšovania procesov a realizuje sa prostredníctvom analýzy procesov. Táto analýza zahŕňa identifikovanie procesov, ich zobrazenie vo vzájomnej súvislosti, umožňuje nájsť nedostatky v procesoch, či efektívne riešenia na ich odstránenie [1].

V tomto článku bude interpretovaný postup a výsledky procesnej analýzy uskutočnenej v malom výrobnom podniku, ktorého hlavná podnikateľská činnosť je smerovaná na výrobu eurookien a dverí.

## II. MAPOVANIE PROCESOV

Nástrojom na mapovanie procesov v podniku je procesná mapa, ktorá graficky znázorňuje procesy v podniku a jednotlivé kroky procesov. Umožňuje lepšie pochopiť procesy a zistiť, ktoré procesy prinášajú pre zákazníka úžitok [1].

Identifikácia procesov výrobného podniku bola uskutočnená po získaní potrebných informácií a následne bola procesná mapa namodelovaná v prostredí nástroja QPR ProcesGuide.



Obr. 1 Procesná mapa

Vo výrobnom podniku boli identifikované dva hlavné procesy, výroba a montáž, a niekoľko riadiacich a podporných procesov.

### III. MODELOVANIE A SIMULÁCIA PROCESOV

Modelovanie a simulácia procesov vybraného podniku boli realizované v nástroji QPR ProcessGuide.

#### A. Modelovanie

Výsledkom modelovania je procesný model, ktorý zachytáva všetky činnosti potrebné pre uskutočnenie vybraných procesov. Väčšina nástrojov, ktoré sa využívajú na modelovanie podnikových procesov sú založené na grafickom štandarde BPMN, ktorý umožňuje ľahko pochopiť modelované procesy všetkým užívateľom [1].

Procesný model výrobného podniku pozostáva z činností uskutočňovaných v rámci hlavných procesov, čiže procesov výroby a montáže. Tieto procesy sú v procesnom modeli dekomponované na jednotlivé podprocesy takto:

#### Podprocesy výroby:

- zameranie okien,
- nákup dreva,
- úvodná výroba,
- príprava,
- nákup kovania,
- nákup skla,
- povrchová úprava,
- kompletizácia.

#### Podprocesy montáže:

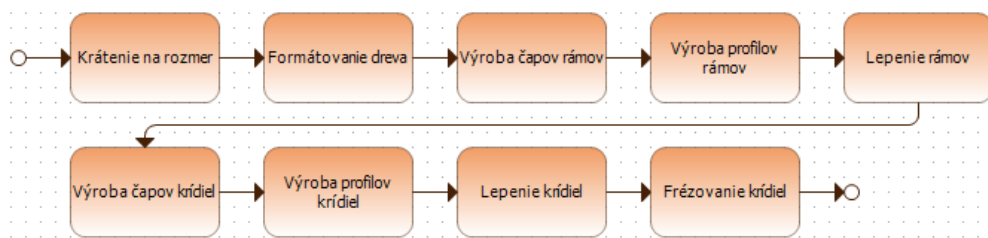
- hlavná montáž,
- ukončovacie práce.

Procesný model podrobne zachytáva všetky činnosti, ktoré sú potrebné pre uskutočnenie každého z uvedených podprocesov. Okrem toho, obsahuje tiež informácie o zdrojoch, potrebných pre vykonanie jednotlivých činností.

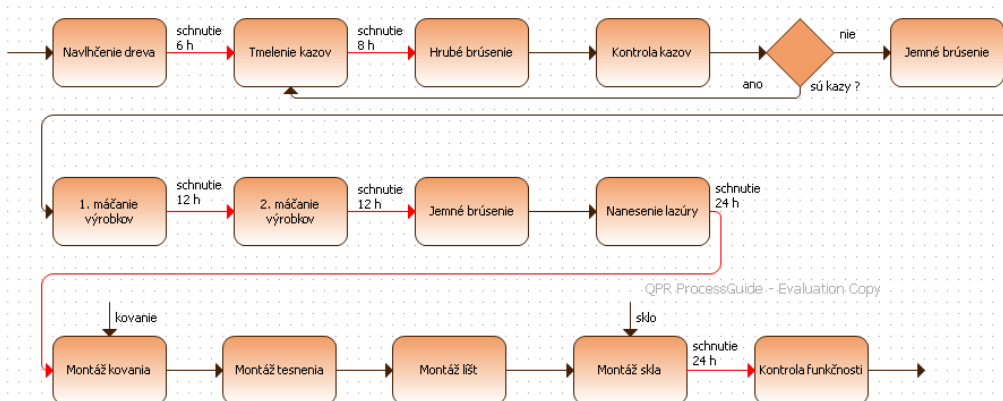
#### B. Simulácia

Simulácia umožňuje namodelované procesy poznať, upravovať a navrhovať nové procesy. Pomáha hodnotiť súčasné procesy, porovnávať ich s novými, či upravenými procesmi a na základe zistení odhadnúť vplyv, ktorý majú navrhované zmeny. Výhodou simulácie je možnosť overiť výsledky navrhovaných zmien ešte pred zavedením týchto zmien do praxe [1].

V prípade simulovania procesného modelu výrobnej firmy boli na simuláciu vybrané len určité časti procesu. Výber týchto častí procesu bol uskutočnený po definovaní cieľa simulácie, ktorým je skrátenie času potrebného pre výrobu a optimálnejšie využitie zdrojov. Simulované boli podprocesy, ktoré sú priamo vykonávané vo výrobnej hale a patria do procesu Výroba, pričom ide celkovo o štyri podprocesy rozdelené na dve simulované časti.



Obr. 2 Prvá simulovaná časť



Obr. 3 Druhá simulovaná časť

#### Navrhované riešenie pre zlepšenie procesov

Po nastavení simulácií a prehodnotení ich výsledkov bol predložený návrh pre zlepšenie procesov. Navrhovaným riešením pre zlepšenie procesov vo výrobnom podniku je zmena prevádzky z jednozmennej na dvojzmennú. V podniku doposiaľ pracovali štyria výrobní pracovníci pri jednozmennej prevádzke. Simuláciou sa potvrdila úspešnosť navrhovaného riešenia, ktoré prináša zmenu prevádzky na rannú v trvaní od 6:00 do 14:00 hod. a odpoľudňajšiu v trvaní od 14:00 do 22:00 hod.. Na každej zmene by pracovali vždy dvaja výrobní pracovníci.

#### Vyhodnotenie navrhnutého riešenia

Uskutočnenými simuláciami sa preukázala úspešnosť navrhovaného riešenia, ktorými sa naplnili stanovené ciele simulácie.

Tab. 1 Porovnanie jednozmennej a dvojzmennej prevádzky

Zmennosť prevádzky	Počet pracovníkov na zmene	Čas potrebný pre výrobu vo výrobnej hale	Vplyv zmienosti na výrobný proces
1-zmenná	4 pracovníci	10 dní a 1,5 h	- Čakacie doby pri paralelnej výrobe. - Neefektívne využitie pracovníkov. - Dlhý čas výroby.
2-zmenná	2 pracovníci	7 dní a 14 h	+ Optimálnejšie využitie strojov. + Efektívnejšie využitie pracovníkov. + Skrátenie času výroby. + Flexibilnejšia montáž.



#### IV. VYHODNOTENIE VÝKONNOSTI PODNIKOVÝCH PROCESOV

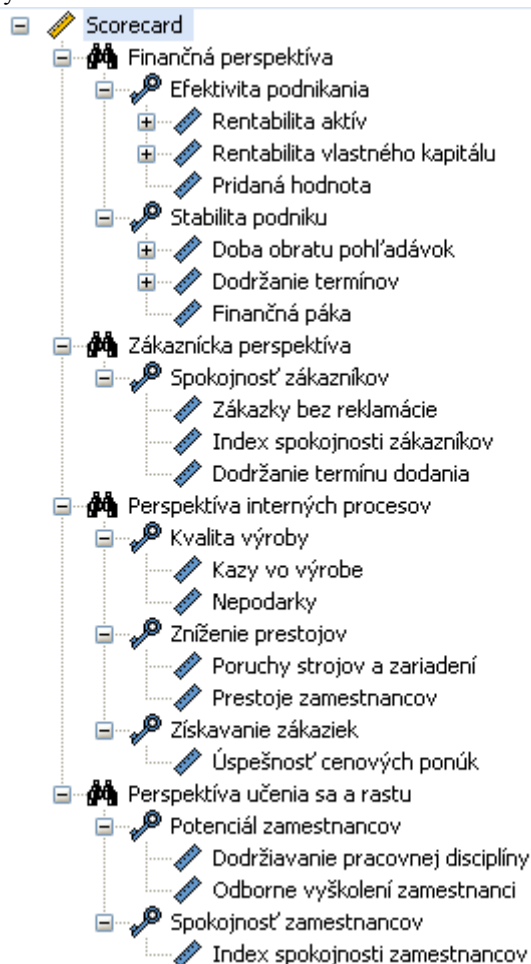
Vyhodnotenie výkonnosti procesov v podniku sa uskutočňuje na základe merania tejto výkonnosti, pričom základným prvkom pre meranie je stanovenie kľúčových ukazovateľov výkonnosti (KPI). V súčasnosti spájajú systémy na meranie výkonnosti finančné aj nefinančné ukazovatele, ktoré sa usilujú o to, aby nedochádzalo k zlepšovaniu určitej oblasti na úkor iných. Žiaden z týchto systémov nie je univerzálny, a preto je dôležité zvoliť správnu metódu a súbor ukazovateľov vhodných na meranie výkonnosti [1].

##### *Balanced Scorecard*

Tento systém na meranie, vyhodnocovanie a riadenie výkonnosti je založený na vyvážení finančných a nefinančných ukazovateľov a zameriava sa na dosahovanie výsledkov v budúcnosti, ktoré sú odvodené od stratégie a vízie [1]. Zameriava sa na štyri oblasti – perspektívy podnikateľskej činnosti a to na finančnú perspektívu, zákaznícku perspektívu, perspektívu interných procesov a perspektívu učenia sa a rastu [1].

##### *Postup vyhodnotenia výkonnosti procesov*

Na vyhodnotenie výkonnosti procesov vo výrobnom podniku metodikou BSC bol využitý nástroj QPR ScoreCard. Prvým krokom pri realizácii Balanced Scorecard (BSC) bolo zistiť podnikovú víziu a stratégiu a na základe nich stanoviť v rámci každej perspektívy ciele, ktorých dosahovanie sa bude sledovať. Každé mu cieľu, čiže kritickému faktoru úspechu (CSF), boli priradené meradlá – kľúčové indikátory výkonnosti (KPI). Tieto indikátory slúžia na meranie dosahovania stanovených cieľov.



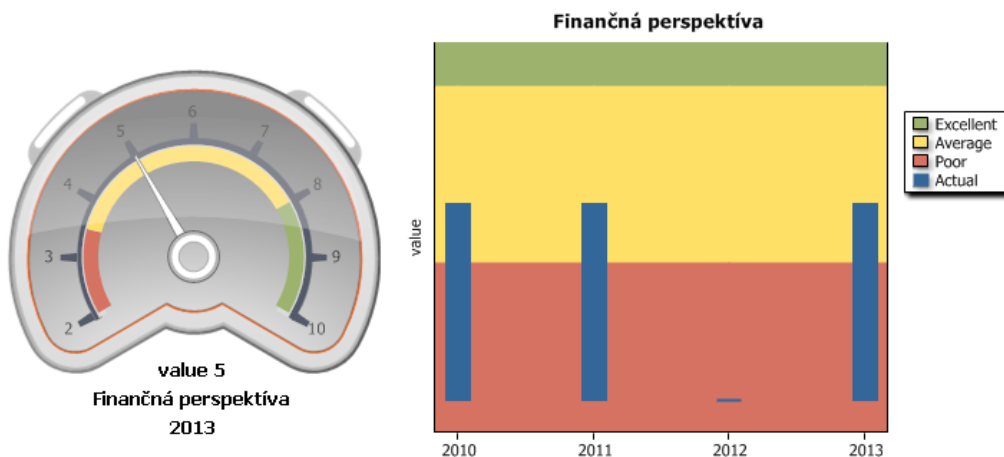
h

**Obr. 4 Zadané perspektívy, CSF a KPI**

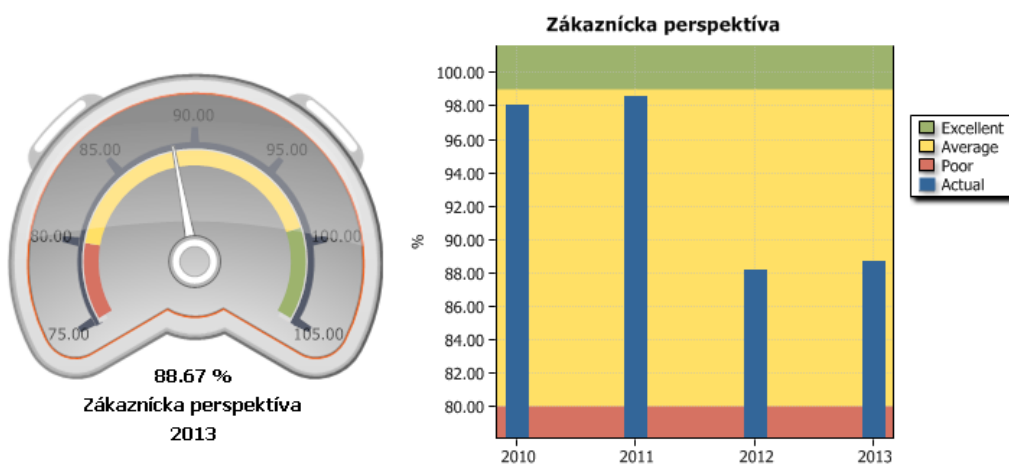
Následne sa v softvéri zadali potrebné nastavenia a hodnoty jednotlivým meradlám. Vyhodnotenie výkonnosti v rámci cieľov a perspektív bolo automaticky vypočítané na základe zadaných funkcií a pomocou hodnôt príslušných meradiel.

##### *Zhodnotenie výsledkov*

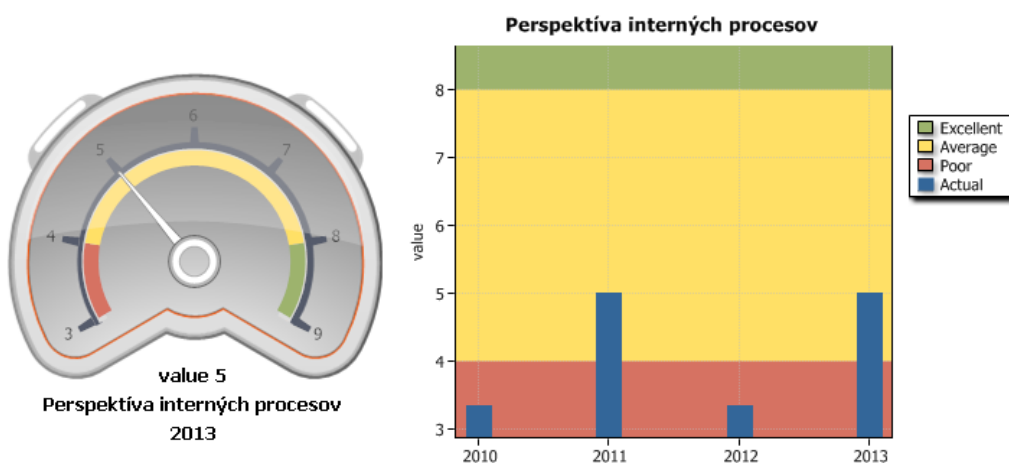
Výkonnosť podnikových procesov sa sledovala v rámci jednotlivých perspektív a priniesla nasledujúce výsledky:



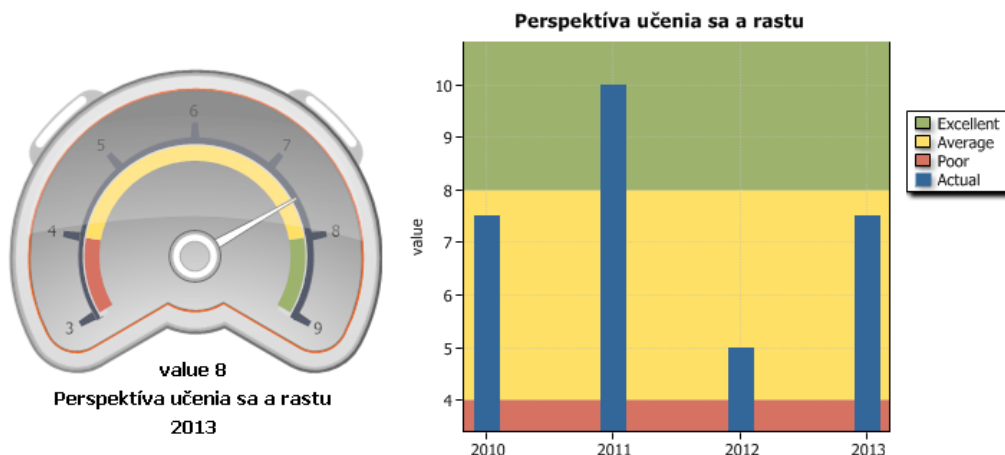
Obr. 5 Finančná perspektíva



Obr. 6 Zákaznícka perspektíva



Obr. 7 Perspektíva interných procesov



Obr. 8 Perspektíva učenia sa a rastu

Zhodnotením zistených výsledkov môžeme poznamenať, že podnik je značne zadlžený, no táto zadlženosť sa každým rokom znižuje a pri udržaní takéhoto trendu si firma môže v budúcnosti zlepšiť svoju finančnú situáciu. Výsledky zákazníckej perspektívy a perspektívy interných procesov poukazujú na to, že navrhované riešenie pre zlepšenie procesov by malo mať kladný vplyv na oblasť spokojnosti zákazníkov aj interné procesy vo firme.

#### V. ZÁVER

Článok sa zaoberal problematikou procesnej analýzy a jej prínosom pre malé podniky. Analýza procesov bola realizovaná v malom výrobnom podniku a zahŕňala identifikovanie procesov a zostrojenie procesnej mapy, namodelovanie hlavných procesov, simulovanie vybraných podprocesov a návrh riešenia pre zlepšenie procesov v podniku. Následne bola meraná a vyhodnotená výkonnosť podnikových procesov metodikou Balanced ScoreCard. Výsledky hodnotenia výkonnosti procesov poukazujú na možný pozitívny vplyv na výkonnosť podniku a to implementáciou navrhnutého riešenia pre zlepšenie procesov.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja (60%). Podporujeme výskumné aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ. Príspevok bol tiež podporený grantom z KEGA - 021TUKE-4/2012 (40%).

#### REFERENCIE

- [1] KONKOLOVÁ, Gabriela: *Procesná analýza výrobnjej firmy s využitím vybraných technológií*. Diplomová práca. Košice: KKUI FEI TUKE, 2014. 119 s.
- [2] Managementmania: *Analýza procesov (procesná analýza)* [online]. Posledná aktualizácia: 17.06.2013. [cit. 2013-08-01]. Dostupné na internete:<<https://managementmania.com/sk/analiza-procesov-procesne-analyza>>.
- [3] VULLERS, M., H., Jansen – NETJES, M.: *Business Process Simulation - A Tool Survey* [online]. [cit. 2013-09-12]. Dostupné na internete:<<http://sistemas-humano-computacionais.wdfiles.com/local--files/capitulo%3Acolaboracao-auxiliada-por-computador/%5BJansen-Vullers%202006%5D%20Business%20Process%20Simulation%20-%20A%20Tool%20Survey.pdf>>.
- [4] GAVUROVÁ, Beáta: *Meranie výkonnosti v organizáciách s dôrazom na aplikáciu systému Balanced Scorecard*. Košice: TUKE, Ekonomická fakulta, 2010. ISBN 978-80-553-0437-3.
- [5] KAPLAN, S., Robert – NORTON, P., David: *The Balanced Scorecard: translating strategy into action*. Boston: Harvard Business School Press, 1996. 322 s. ISBN 0-87584-651-3.

# Tvorba reportovacej vrstvy v prostredí Microsoft SQL Server 2012

*Daniela JUDIČÁKOVÁ, Peter MICHALIK, Iveta ZOLOTOVÁ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

daniela.judicakova@student.tuke.sk, peter.michalik.2@tuke.sk,  
iveta.zolotova@tuke.sk

**Abstrakt** — Článok sa zaoberá tvorbou experimentálnej reportovacej vrstvy v prostredí Microsoft SQL Server 2012. Stručne opisuje Business Intelligence a vysvetľuje reporting. Článok ďalej stručne charakterizuje prostredie MS SQL Server 2012 a zameriava sa na oblasť reportovacích služieb. Cieľom je v tomto prostredí vytvoriť reportovacia vrstvu z reálnych dát, pričom rozoberá postup v reportovacích nástrojoch Microsoft Visual Studio a Report Builder, kde je podrobne priblížený návrh reportov a analýza dát pomocou týchto nástrojov.

**Kľúčové slová** — Business Intelligence, Reporting, MS SQL Server 2012

## I. ÚVOD

Súčasný Business Intelligence je neodmysliteľnou súčasťou podnikania vo vzájomne prepojenom svete riadenom informáciami. Rýchlejšie spracovanie zozbieraných údajov, ich analýza a následná interpretácia napomáha firmám pri rozhodovaní. Bez možnosti analýzy a efektívneho zobrazenia výsledkov je množstvo získaných dát doslova zbytočných. Spoločnosť Microsoft reagovala na dopyt systémom SQL Server 2012, ktorý ponúka komplexné funkcie BI a taktiež umožňuje prezentovať dáta v zrozumiteľnej podobe.

## II. BUSINESS INTELLIGENCE A REPORTING

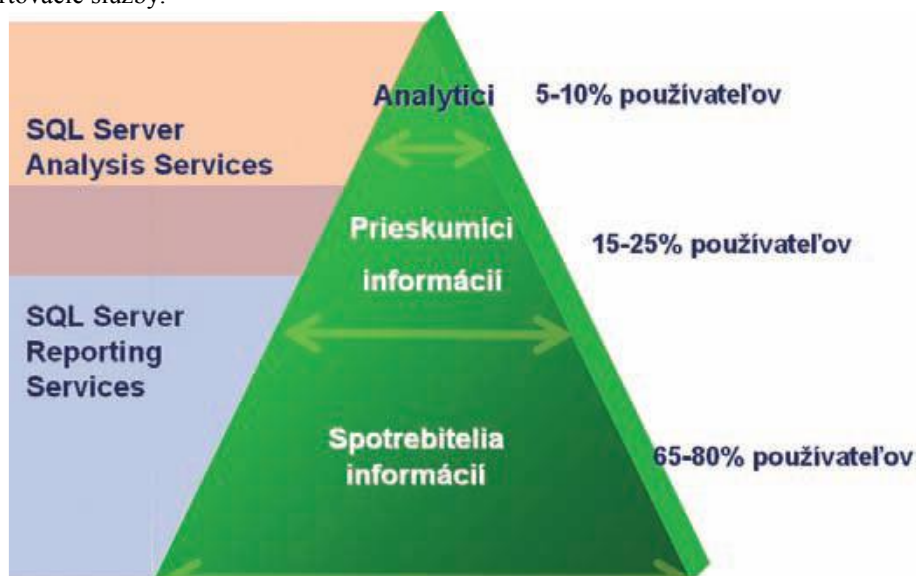
Business Intelligence sú systémy pre podporu rozhodovania. Prostredníctvom Business Intelligence sa vykonáva zber údajov, ich dôkladná analýza a zrozumiteľná interpretácia, teda vytváranie reportov, ktoré napomáhajú plánovaniu a rozhodovaniu vo firmách. Úlohou Business Intelligence je na požiadanie poskytovať informácie v reálnom čase a zobrazovať ich v zrozumiteľnej forme. BI poskytuje používateľom nástroj ako sa rýchlo dostať k požadovaným informáciám. Vďaka BI sa dokážu manažéri rýchlo rozhodovať. Keďže majú ľahký prístup k informáciám a možnosť ich interaktívne prehľadávať a porovnávať, pomáha im to prijímať správne strategické rozhodnutia. Zmyslom BI je získať jasný a zrozumiteľný prehľad zozbieraných informácií, ktorý bude ľahko prístupný v prehľadnej forme.

Manažéri sa potrebujú rýchle a správne rozhodnúť bez toho, aby poznali celý proces Business Intelligence. Dôležitý je pre nich až validný výsledok. Reportingom teda možno nazvať posledný stupienok celého procesu BI. Prostredníctvom reportingu získa užívateľ informácie vo forme, ktorá je mu graficky príjemná, funkcionálne dostatočná a jednoduchá na obsluhu. To je tým pravým dôvodom, prečo je celý BI proces dovŕšený práve fázou reportingu. Behom procesov sa dáta extrahujú, transformujú, ukladajú, prechádzajú krokmi data miningu, ale bez jasne a účelne zobrazeného výstupu vo forme prehľadnej tabuľky či grafu, v ktorom sa cieľový užívateľ nestratí, by to nemalo žiaden význam. Úlohou reportovacích služieb je teda včas a vo vhodnej forme poskytovať podklady pre podporu rozhodovania na všetkých stupňoch organizačnej štruktúry podniku [1] [2].

### III. MICROSOFT SQL SERVER 2012

MS SQL Server 2012 je okrem iného aj nástroj pre vytváranie reportov. Reportovacie služby sú služby integrované do SQL Serveru 2012 a nasadzujú sa v rámci podnikovej informačnej infraštruktúry, aby užívateľom pomáhali efektívne pristupovať k údajom a manažérom pomáhali v procese rozhodovania [1].

Podľa prieskumov bolo zistené, že najväčšia skupina užívateľov používa informácie vo forme rôznych výpisov a reportov. Prístup k údajom vo forme reportov užívateľom zabezpečujú reportovacie služby.



Obr. 1 Percentuálne využitie jednotlivých BI technológií [3]

Prostredníctvom služby *Reporting Services* je možné vytvárať reporty, spravovať ich a sprístupňovať užívateľom. Reporty môžu byť vytvárané priamo nad relačným databázovým severom alebo nad databázami a kockami *Analysis Services*. Návrh reportu je možné realizovať prostredníctvom rôznych nástrojov v prostredí *BI Visual Studio 2010*, napr. *Report Builder*. Výsledné reporty môžu byť generované buď na požiadavku alebo v opakujúcich sa intervaloch. Môžu nadobúdať rôzne formáty, napríklad v podobe HTML stránok, formátu PDF, aplikácie zo skupiny *MS Office*, najčastejšie *Microsoft Excel*, čistého textu alebo aj ako obrázkov [4].

### IV. PRÁCA V PROSTREDÍ MS SQL SERVER 2012

Tvorba reportovacej vrstvy bola realizovaná na reálnych dátach firmy. Použité boli programy *SQL Server 2012 Management Studio*, *Microsoft Visual Studio* a *Report Builder*.

#### A. Vytvorenie a naplnenie relačnej databázy

Program *SQL Server 2012 Management Studio* slúži na vytvorenie databázy, a taktiež na vytváranie, editovanie a napĺňanie tabuliek v databáze.

Výsledkom je databáza naplnená dátami, ktorá bude potrebná a využívaná pri práci s nástrojmi na tvorbu reportov.

#### B. Práca v programe *MS Visual Studio*

Práca je rozdelená na návrh reportu a analýzu dát.

##### *Návrh reportu*

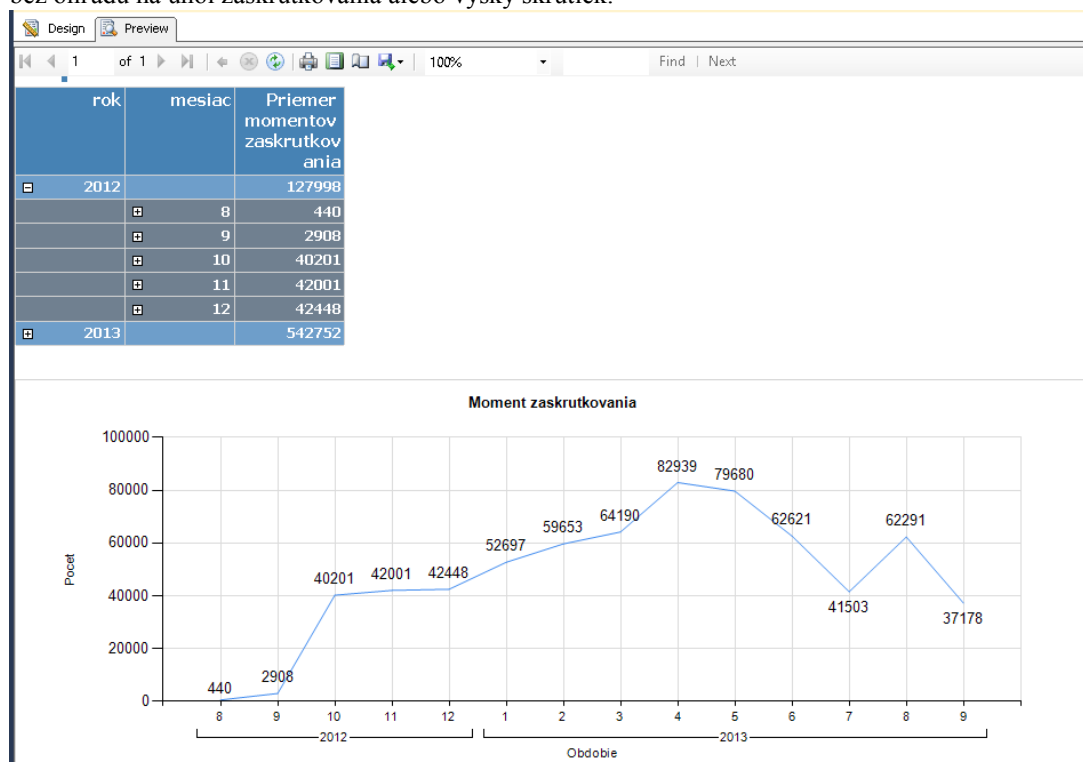
Postup návrhu reportu pozostáva z niekoľkých na seba nadväzujúcich krokov:

- vytvorenie dátového zdroja,
- návrh SQL kódu pre výber požadovanej množiny údajov,
- výber typu tabuľky,

- návrh dizajnu tabuľky (rozmiestnenia údajov),
- výber štýlu reportu, grafická úprava.

### Analýza dát

Úlohou bolo nájsť obdobie, v ktorom bolo vykonaných najviac momentov zaskrutkovania, bez ohľadu na uhol zaskrutkovania alebo výšky skrutiek.



Obr. 2 Analýza dát podľa obdobia

Vytvorená bola tabuľka, kde sa pomocou ikoniek so znamienkom (+) možno vnárať na nižšie úrovne a podrobne pozrieť počty v konkrétnom mesiaci. Taktiež bola vytvorená krivka, na ktorej vidieť, v ktorom období počet momentov rastie, a v ktorom naopak klesá. Najvyšší počet momentov je zaznamenaný v apríli roku 2013.

### C. Práca v programe Report Builder

Klientska aplikácia Report Builder je určená pre užívateľov, ktorí si chcú reporty vytvárať sami, teda ad hoc reporty.

#### Návrh reportu

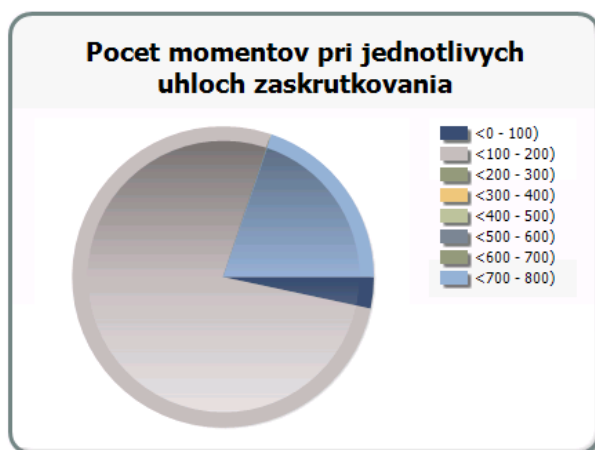
Postup návrhu reportu pozostáva z niekoľkých na seba nadväzujúcich krokov:

- vytvorenie dátového zdroja,
- návrh SQL kódu pre výber požadovanej množiny údajov,
- návrh rozmiestnenia údajov v tabuľke,
- návrh dizajnu tabuľky,
- výber štýlu reportu, grafická úprava.

### Analýza dát

Prvou úlohou bolo nájsť taký interval pri atribúte uhlo zaskrutkovania, v ktorom bolo vykonaných najviac momentov zaskrutkovania.

UHOL ZASKRUT Priemer	Priemer momentov zaskrutkovania
<0 - 100)	22590
<100 - 200)	514844
<200 - 300)	94
<300 - 400)	82
<400 - 500)	13
<500 - 600)	10
<600 - 700)	340
<700 - 800)	132777
<b>Total</b>	<b>670750</b>



Obr. 3 Analýza dát na základe uhlu zaskrutkovania

Pri tejto analýze je vidieť, že najvyšší počet momentov sa vykazuje pri uhle <100 – 200). Pri tomto intervale je vykonaných až 514 844 momentov zaskrutkovania z celkového počtu 670 750, čo vykazuje viac ako trištvrtinu z celkového počtu.

Ďalšou úlohou bolo pri vybranom uhle a pri vybranej výške skrutky po skrutkovaní ukázať počet momentov za sledované časové obdobie.

UHOL ZASKRUT Priemer1	VÝSKA PO SKRUT Priemer1	2012										2013 Total	
		Total	1	2	3	4	5	6	7	8	9	Total	Total
<100 - 200)	<0 - 1)	79146	39961	41995	47994	66576	53491	47616	36621	40136	25447	399837	478983
	Total	79146	39961	41995	47994	66576	53491	47616	36621	40136	25447	399837	478983
Total		79146	39961	41995	47994	66576	53491	47616	36621	40136	25447	399837	478983



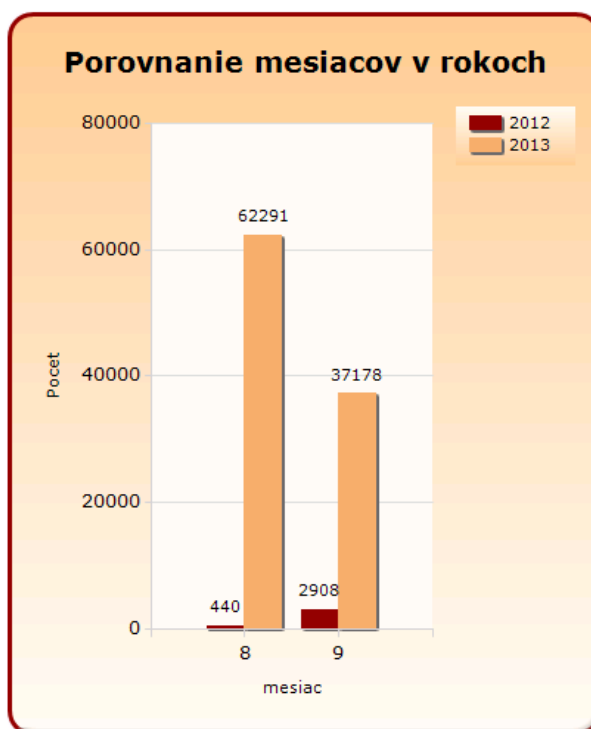
Obr. 4 Analýza dát na základe vybraného uhla a výšky skrutky po skrutkovaní

Analýza ukázala, že počet momentov od októbra 2012 do apríla 2013 pravidelne stúpala a od tohto momentu začal v nasledujúcich mesiacoch klesať.

V nasledujúcej úlohe bolo cieľom porovnať počty momentov pri rovnakom mesiaci, ale

v inom roku.

rok1	mesiac1	Priemer momentov zaskrutkova nia
2012	8	440
	9	2908
	<b>Total</b>	<b>3348</b>
2013	8	62291
	9	37178
	<b>Total</b>	<b>99469</b>
<b>Total</b>		<b>102817</b>



Obr. 5 Porovnanie dát pri rovnakých mesiacoch v inom roku

Výsledkom analýzy je porovnanie mesiacov augusta a september v rokoch 2012 a 2013. Jasne vidieť, že firma v roku 2013 podstatne zvýšila počet momentov zaskrutkovania oproti predchádzajúcemu roku.

## V. ZÁVER

Článok sa zaoberal problematikou reportovania a prácou s prostredím MS SQL Server 2012. Poznatky týkajúce sa pojmov Business Intelligence a reportovania boli využité pri práci s vybraným prostredím a pri následnej analýze dát. Analýza bola realizovaná na reálnej vzorke dát a celková práca zahŕňala úpravu dát, tvorbu relačnej databázy, import dát, návrh reportov a výslednú analýzu dát.

## POĎAKOVANIE

Táto práca bola podporená grantom KEGA - 021TUKE-4/2012(70%) a grantom VEGA - 1/0286/11 (30%).

## REFERENCIE

- [1] JUDIČÁKOVÁ, Daniela: Tvorba reportovacej vrstvy v prostredí Microsoft SQL Server 2012. Diplomová práca. Košice: KKUI FEI TUKE, 2014. 62 s.
- [2] PETRÁŠEK, Michael: Business intelligence a reporting nejen pro krizové období [online]. systemonline.cz. 2009. ISSN: 1802-615X. [cit. 2013-11-24]. Dostupné na internete: <<http://www.systemonline.cz/business-intelligence/business-intelligence-a-reporting-nejen-pro-krizove-obdobi.htm>>.
- [3] LACKO, Luboslav: Business Intelligence v SQL Serveri 2005 [online]. Praha, Microsoft s.r.o., 2005. [cit. 2014-4-20]. Dostupné na internete: <[http://download.microsoft.com/download/1/B/0/1B055693-B82B-491F-9FD6-4FE9B17C77DF/Zbornik\\_BI\\_SQL2005.pdf](http://download.microsoft.com/download/1/B/0/1B055693-B82B-491F-9FD6-4FE9B17C77DF/Zbornik_BI_SQL2005.pdf)>.
- [4] ČIPKALA, Jaroslav: Analýza údajov z AIS2 s využitím nástrojov Business Intelligence: Diplomová práca [online]. Banská Bystrica, UMB, EF, 2013. 73 s. [cit. 2014-4-18]. Dostupné na internete: <<http://diplomovka.sme.sk/zdroj/3778.pdf>>.



# Úlohy objavovania znalostí v prostredí technológie Microsoft SQL Server 2012

<sup>1</sup>Katarína JEŇOVÁ, <sup>2</sup>Peter MICHALIK, <sup>3</sup>Iveta ZOLOTOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

<sup>1</sup>katarina.jenova@student.tuke.sk, <sup>2</sup>peter.michalik.2@tuke.sk, <sup>3</sup>iveta.zolotova@tuke.sk

**Abstrakt** — Článok sa zaoberá Data Miningom a využitím jeho úloh pri návrhoch modelov pomocou technológie Microsoft SQL Server 2012. Na poukázanie vlastností a možnosti využitia tohto nástroja boli vytvorené modely dolovania dát, realizované na dátach z výrobnéj linky pre kompletizáciu motora.

**Kľúčové slová** — Data Mining, Microsoft SQL Server 2012, model, algoritmus

## I. ÚVOD

Akýkoľvek záznam, či už numerický alebo alfanumerický, predstavuje údaj alebo dáta. Údajom je pritom každá správa bez ohľadu, či má pre nás nejaký informačný obsah alebo nie. Každá firma, ktorá o svojej aktivite uchováva dáta predpokladá, že tieto údaje budú mať pre ňu nejakú výpovednú hodnotu. Avšak nie všetky prospešné informácie sú na prvý pohľad zreteľné. Na prekonanie akejsi bariéry neviditeľných súvislostí slúži dolovanie dát. Aj keď v zahraničí je Data mining už celkom zabehnutou súčasťou podnikov, na Slovensku je využívanie tohto procesu skôr niekde na začiatku.

Cieľom tohto článku bude interpretovať postup aplikovania metódy dolovania znalostí na reálnych dátach a zároveň predstaviť Microsoft SQL Server 2012 ako jeden z nástrojov, ktorý je možno využiť na konštruovanie modelov.

## II. DATA MINING

Data mining je momentálne najrýchlejšie rastúcim segmentom Business Intelligence a v súčasnosti majú túto technológiu implementované všetky významné komerčne dodávané databázové servery.

Dolovanie dát nazývame proces, kedy sa pokúšame za pomoci rôznych matematických metód nájsť vnútorné, zatiaľ nepoznané, súvislosti. Tieto objavené znalosti môžeme neskôr využiť pri prediktívnej analýze [1].

Na základe problému, ktorý je potrebné riešiť, je možné rozhodnúť sa pre niektorý z typov úloh dolovania dát. Medzi najčastejšie používané typy patrí klasifikácia, predikcia, asociačné pravidlá a zhľukovanie. Na splnenie týchto úloh sa využívajú rôzne algoritmy, pričom niektoré algoritmy môžu poskytnúť viac ako len jeden výsledok. K najznámejším algoritmom patrí rozhodovací strom, neurónové siete, asociačné pravidlá, rôzne typy regresii, naive bayes a mnoho ďalších [2].

### III. TVORBA MODELOV

Experimentálne úlohy boli vytvárané na reálnych dátach spoločnosti B/S/H/ Drives and Pumps s.r.o. Michalovce, pomocou nástroja Microsoft SQL Server 2012 [3].

#### A. Príprava

Použité dáta boli získané z databázy zaznamenajúcej údaje z výrobných linky pre kompletizáciu motora. Sledované obdobie je od augusta 2012 do konca septembra 2013, pričom v tomto období bolo nameraných 670 750 záznamov.

Tieto údaje pozostávali zo 17 atribútov:

- Dátum
- Moment zaskrutkovania skrutiek 1 – 4
- Uhol zaskrutkovania skrutiek 1 – 4
- Výška skrutky pred zaskrutkovaním 1 – 4
- Výška skrutky po zaskrutkovaní 1 – 4

Tieto dáta sa následne upravili, spracovali a rozdelili do tabuliek tak aby bolo možné s nimi ďalej pracovať a využívať pri tvorbe multidimenzionálnej kocky ako aj pri tvorbe data miningových modelov.

Ďalším z krokov bolo vytvorenie databázy prostredníctvom administratívneho nástroja SQL Server Management Studio. V databáze sa následne vytvárali databázové tabuľky, s ktorými sa v úlohách pracovalo. Tieto tabuľky reprezentujú dimenzie a tabuľka faktov. Vytvorené boli na základe DDL príkazu, ktorý sa vložil do pracovného priestoru databázy. Tieto tabuľky sú však ešte prázdne a teda bolo potrebné ich naplniť údajmi, ktoré boli predspracované už v prvom kroku prípravy, pričom každá z tabuliek sa bude naplňať osobitne.

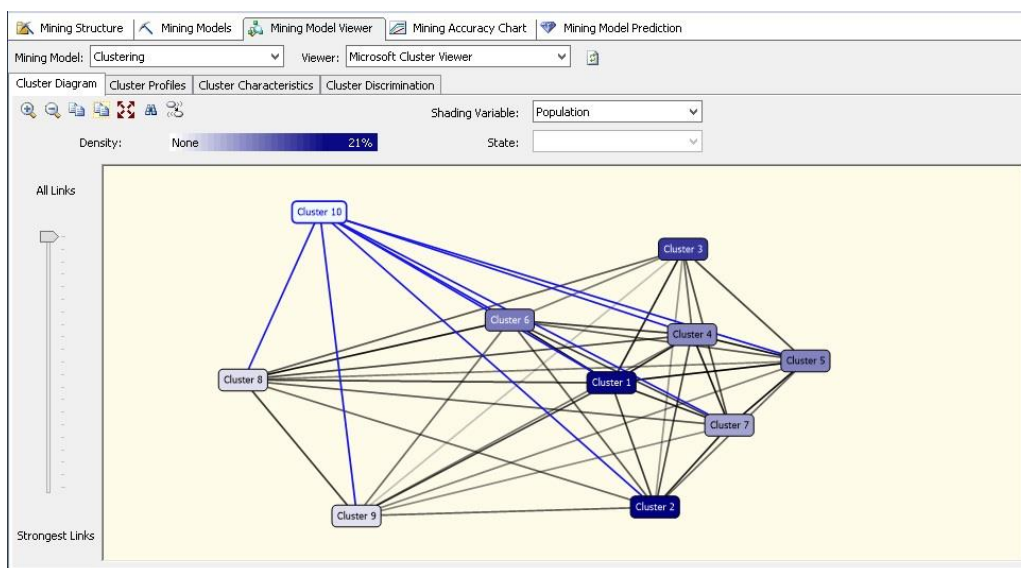
S takto pripravenou databázou bolo možné začať so samotnou tvorbou modelov.

#### B. Tvorba modelov

Technológia Microsoft SQL Server 2012 ponúka sedem algoritmov pre tvorbu modelov. Ako príklad je možné uviesť niektoré z možných algoritmov na ktorých bolo uplatnené dolovanie dát, a to zhľukovanie, Naive Bayes a asociačné pravidlá.

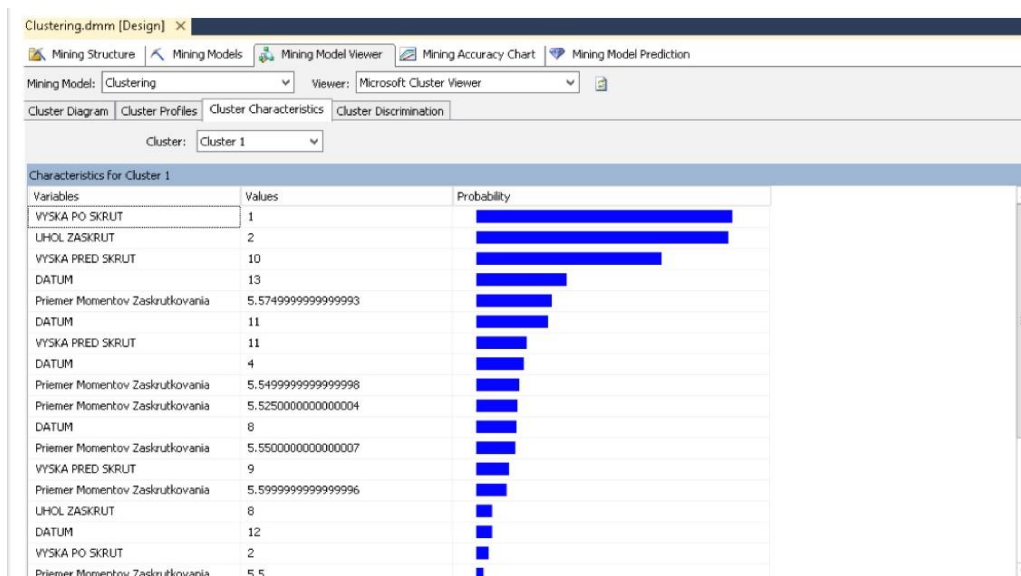
#### Zhľukovanie

Zámerom zhľukovania je nájsť také zhľuky, ktorých dáta majú podobné vlastnosti. Pomocou dialógového sprievodcu sa nastavujú vstupné parametre a požadovaný výstup, ktoré sú potrebné pre vytvorenie zhľukov. Po vytvorení a spustení modelu sa na pracovnej ploche zobrazia všetky zhľuky poprepájané tieňovanými čiarami, ktoré vyjadrujú silu vzťahom medzi nimi.



Obr. 1 Model zhľukovania

Tak tiež jednou s možnosťami, ktoré tento nástroj ponúka, je prezretie si vlastností konkrétneho zhluku. Pre každý zhluk sú vypísané atribúty spolu s ich hodnotami a pravdepodobnosťou.



Obr. 2 Vlastnosti zhluku č.1

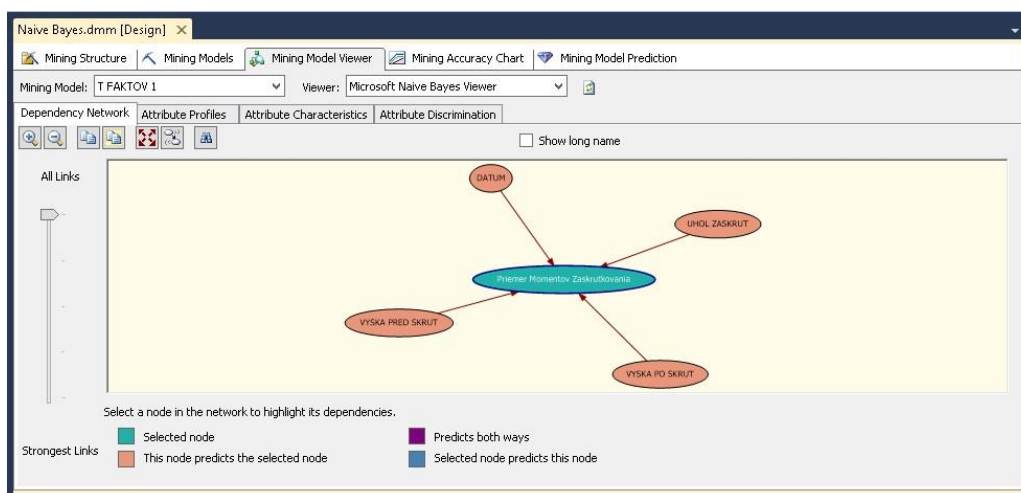
*Výsledok*

Pre zhluk číslo jeden boli zistené nasledujúce atribúty a ich pravdepodobnosti výskytu pre daný zhluk:

Atribút	Hodnota	Percento
VYSKA PO SKRUT	1	98.105%
UHOL ZASKRUT	2	93.959%
VYSKA PRED SKRUT	1	68.824%
DATUM	13	33.812%
Priemer Momentov Zaskrutkovania	5.574999	28.218%

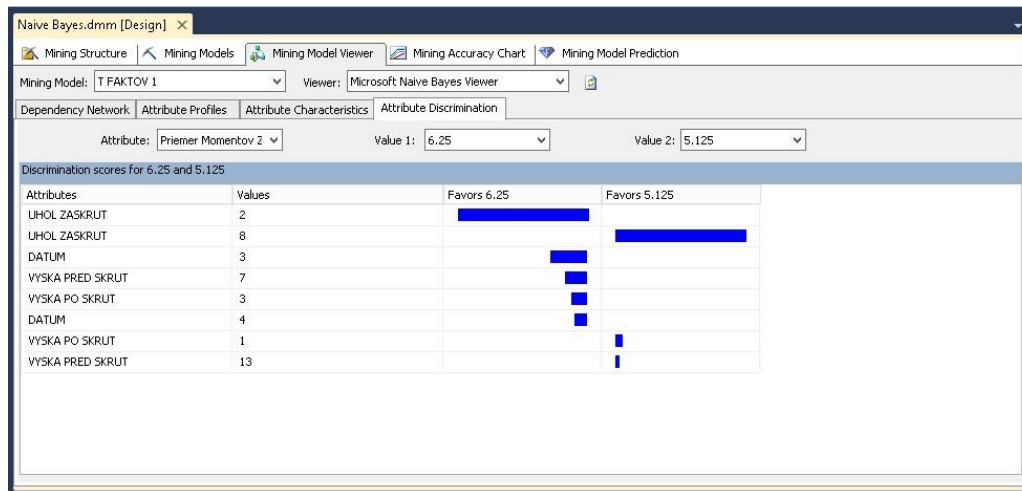
*Naive Bayes*

Úlohou modelu Microsoft Naive Bayes bolo určiť vplyv jednotlivých atribútov na zvolené hodnoty Priemerov momentov zaskrutkovania.



Obr. 3 Model Naive Bayes

Pre zistenie vplyvu jednotlivých atribútov na vybranú hodnotu sa zvolí záložka Attribute Discrimination, ktorú ponúka tento nástroj. Pre zvolený atribút Priemer momentov zaskrutkovania sa vyberie napríklad **Hodnota 1: 6,25** a **Hodnota 2: 5,125**.

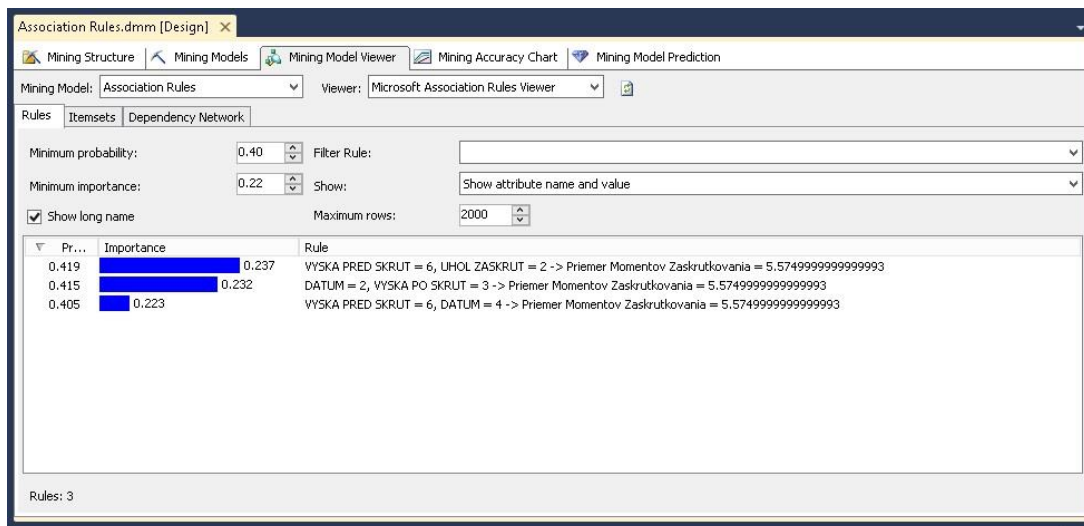


**Obr. 4** Vybrané hodnoty pre zistenie vplyvu

Na základe zvoleného modelu možno povedať, že atribút UHOL ZASKRUT pri hodnote = 2, vplýva 100% na priazeň hodnoty 6,25, zatiaľ čo ten istý atribút len pri hodnote = 8, vplýva 100% na druhú, nami zvolenú hodnotu – 5,125. Atribúty DATUM (hodnota = 3 a 4), VYSKA PRED SKRUT (hodnota = 7) a VYSKA PO SKRUT (hodnota = 3) vplývajú na hodnotu 6,25 s menej než 30 percentami. Taktiež atribúty VYSKA PO SKRUT (hodnota =1) a VYSKA PRED SKRUT (hodnota = 13) vplývajú na hodnotu 5,125 zanedbateľným percentom – menším ako 6%.

*Asociačné pravidlá*

Pomocou asociačného pravidla bolo potrebné nájsť také vzory, pri ktorých priemery momentov zaskrutkovania boli pod hodnotou šiestich sekúnd, pričom minimálna pravdepodobnosť je defaultne nastavená na hodnotu 0,40. Po nastavení vstupných a výstupných hodnôt, model objavil tri pravidlá s nasledujúcimi výsledkami.



**Obr. 5** Asociačné pravidlá

### Výsledok

Ak je VYSKA PRED SKRUT = 6 UHOL ZASKRUT = 2, potom Priemer Momentov Zaskrutkovania = 5,57499. Pravdepodobnosť, že taká možnosť nastane je 0,419. Pri tomto pravidle bolo klasifikovaných 32 prípadov, pričom užitočnosť pravidla je 0,237.

Ak je VYSKA PO SKRUT = 3 a DATUM = 2, potom Priemer Momentov Zaskrutkovania = 5,57499. Pravdepodobnosť, že taká možnosť nastane je 0,415. Pri tomto pravidle bolo klasifikovaných 40 prípadov, pričom užitočnosť pravidla je 0,232.

Ak je VYSKA PRED SKRUT = 6 a DATUM = 4, potom Priemer Momentov Zaskrutkovania = 5,57499. Pravdepodobnosť, že taká možnosť nastane je 0,405. Pri tomto pravidle bolo klasifikovaných 32 prípadov, pričom užitočnosť pravidla je 0,223.

### IV. ZÁVER

Článok sa zaoberal využitím technológie Microsoft SQL Server 2012 ako jedným z možných nástrojov na dolovanie dát. Tento nástroj využíva viacero algoritmov, ktoré je možné použiť pre objavovanie rôznych doposiaľ nezistených súvislostí medzi zozbieranými dátami. Tieto algoritmy môžu byť použité pri klasifikačných ako aj pri predikčných úlohách Data Miningu.

### POĎAKOVANIE

Táto práca bola podporená grantom KEGA - 021TUKE-4/2012(70%) a grantom VEGA - 1/0286/11 (30%).

### REFERENCIE

- [1] Ebookbrowse. Dolovani pdf. [online] . [cit. 2013-4-20]. Dostupné na internete: <<http://ebookbrowse.com/dolovani-pdf-d51888469>>
- [2] Zhansheng Zhang. 2010. Application of SQL Server in data mining. [online]. Chinese: Control and Decision Conference (CCDC), 2010. E-ISBN 978-1-4244-5182-1. [cit. 2014-4-27]. Dostupné na internete: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5498995&queryText%3Ddata+mining+sql>>
- [3] JEŇOVÁ, Katarína: *Úlohy objavovania znalostí v prostredí technológie Microsoft SQL Server 2012*. Diplomová práca. Košice: KKUI FEI TUKE, 2014.

# Analýza meteorologických dát s využitím nástroja QlikView

*Pavol HAJASTEK, Peter MICHALIK, Iveta ZOLOTOVÁ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

pavol.hajastek@student.tuke.sk, iveta.zolotova@tuke.sk, iveta.zolotova@tuke.sk

**Abstrakt** — Článok sa zaoberá problematikou Business Intelligence a využitím nástroja QlikView. Opisuje prínosy QlikView a porovnáva tieto prínosy s tradičnými riešeniami Business Intelligence. Hlavným zámerom článku však nie je iba teoretická analýza vlastností a prínosov daného nástroja, ale aj jeho nasadenie, teda implementácia takéhoto informačného systému pre potreby analýzy meteorologických dát vybranej meteorologickej stanice.

**Kľúčové slová** — QlikView, Business Intelligence, Analýza dát, Vizualizácia dát

## I. ÚVOD

Článok popisuje problematiku Business Intelligence, konkrétne nástroja QlikView a oblasti návrhu ako aj vývoja aplikácie v danom nástroji. Načrtnuté prínosy QlikView sú tak aplikované pri tvorbe aplikácie slúžiacej pre vytváranie rôznych meteorologických analýz v rámci možnosti daného nástroja.

## II. QLIKVIEW VERZUS TRADIČNÉ RIEŠENIA BUSINESS INTELLIGENCE

Nástroj QlikView je považovaný za nástroj novej generácie Business Intelligence, takzvaných Data Discovery prístupov k BI a predstavuje akýsi „most“ medzi nedostatkami kancelárskych informačných systémov a „ťažkopádnyimi“, tradičnými riešeniami BI. Konečným používateľom umožňuje skúmanie informácií bez nutnosti využívať preddefinované cesty a to bez potreby zásahov IT-špecialistov [1][2].

Jadrom QlikView je patentovaná in-memory technológia, ktorá predstavuje najväčší rozdiel oproti tradičným riešeniam BI. Za pomoci tejto technológie sú všetky dáta v QlikView načítavané do pamäte RAM, čo umožňuje vykonávať rýchle prepočty aj nad veľkými objemami dát a výsledky vytváraných analýz sú tak zobrazené takmer okamžite. Do tejto pamäte RAM je umožnené v QlikView načítavať dáta z rôznych podnikových systémov, dátových skladoch, dátových súborov, či sociálnych sietí akými sú napríklad Facebook alebo Twitter. Všetky dáta sú navyše vzájomne prepojené na základe názvov stĺpcov jednotlivých tabuliek čo umožňuje skúmanie týchto dát s nekonečnou flexibilitou a takzvaným „asociatívnym“ zážitkom [1][3].

## III. NÁVRH RIEŠENIA

Prvým krokom pre vytvorenie aplikácie pre potreby analýzy meteorologických dát bola kompletná analýza zdrojových dát, požiadaviek na aplikáciu a možnosti v rámci softvérového komponentu QlikView Desktop.

### A. Rozbor zdrojových dát

Zdrojovými dátami pre vytvorenie aplikácie boli dáta meteorologickej stanice SZEŠ Písek, ktoré sú dostupné na internetovej stránke spomínanej meteorologickej stanice. Konkrétne sa jednalo o nasledujúce csv súbory, ktoré obsahujú záznamy vonkajšími senzormi.

- Teplota Out.csv
- Tlak.csv
- Vítr.csv
- Srážky.csv

Ako oddeľovač dát je v týchto súborov použitá čiarka a každý z týchto súborov obsahuje informáciu o dátume a čase merania k jednotlivým záznamom. Komplexným skúmaním týchto dátových súborov boli zistené rôzne chyby, akými sú napríklad prázdne hodnoty, redundantné záznamy, či rôzne anomálie, ktoré museli byť odstránené v ďalšom procese tvorby danej aplikácie.

Keďže QlikView umožňuje načítavanie dát z viacerých zdrojov boli do daného riešenia integrované ďalšie dátové súbory a to:

- Meniny.xlsx
- CompactCalendar2012.xls
- Glossary.txt

Pridanie týchto dátových súborov umožnilo rozšíriť možnosti aplikácie o vyhľadávanie konkrétneho dňa na základe menín, či sviatkov a vytvorenie takzvaného slovníka meteorologických pojmov a výrazov vďaka súboru Glossary.txt.

#### A. Požiadavky na aplikáciu

Jednotlivé požiadavky boli z dôvodu prehľadnosti a uľahčenia procesu tvorby danej aplikácie rozdelené na niekoľko častí a to:

- Výber časového úseku (rok, mesiac, ročné obdobie, deň, časť dňa a čas)
- Zobrazenie základných meteorologických veličín (teplota, vietor, tlak, zrážky, vlhkosť)
  - vývoj priemerných hodnôt
  - vývoj maximálnych a minimálnych hodnôt
  - zobrazenie extrémnych hodnôt
- Zobrazenie pomocných meteorologických veličín
  - pocitová teplota (min/max)
  - rosný bod (min/max)
  - náraz vetra (min/max)
  - prepočítaný tlak na hladinu mora (min/max)
  - kumulované zrážky (min/max)
- Štatistika
  - teplota (min/max)
  - rosný bod (min/max)
  - index tepla (min/max)
  - pocitová teplota (min/max)
  - rýchlosť vetra (min/max)
  - náraz vetra (min/max)
  - tlak (min/max)
  - prepočítaný tlak na hladinu mora (min/max)
  - zrážky (min/max)
  - hodinový úhrn zrážok (min/max)
  - denný úhrn zrážok (min/max)
  - mesačný úhrn zrážok (min/max)
  - vlhkosť (min/max)

#### B. QlikView Desktop

Tvorba aplikácie prebiehala v softvérovom komponente QlikView Desktop určenom pre operačný systém Windows. Tento nástroj je používaný najmä biznis analytikmi, respektíve QlikView developermi na vývoj aplikácií, ktorý zahŕňa tvorbu dátového modelu, SQL výberov, ETL procedúr, či grafického užívateľského prostredia. Umožnené je tiež prehliadanie vytvorených aplikácií a teda tvorba rôznych analýz a podobne.

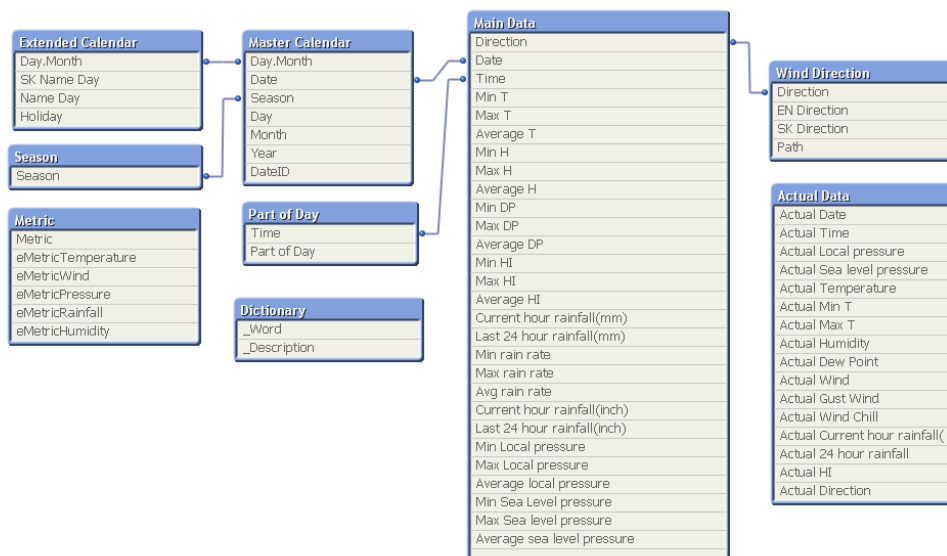
## IV. IMPLEMENTÁCIA RIEŠENIA

Vývoj aplikácie v nástroji QlikView by sme mohli rozdeliť na dve základné časti a to tvorba dátového modelu v prvom kroku a tvorba užívateľského rozhrania pre vytváranie analýz koncovými používateľmi v druhej fáze vývoja.

## A. Dátový model

Dátový model bol konštruovaný za pomoci skriptovacieho jazyka QlikView prostredníctvom ktorého boli jednotlivé tabuľky extrahované z dátových súborov a transformované do potrebnej podoby. Táto transformácia daných dát zahŕňala rôzne agregácie údajov na úrovni jednej hodiny, odstránenie anomálií, prepočty merných jednotiek, výpočty pomocných meteorologických hodnôt na základe stanovených požiadaviek, či pridanie polí s ďalšími potrebnými informáciami pre uskutočňovanie užívateľských výberov.

Vytvorený asociatívny dátový model je vložkovej schémy s jednou tabuľkou faktov, ktorá bola skonštruovaná zlúčením jednotlivých tabuliek obsahujúcich informácie o nameraných hodnotách a niekoľkými ďalšími tabuľkami dimenzií, tak ako môžete vidieť na nasledujúcom obrázku.



Obr. 1 Vytvorený dátový model

## B. Užívateľské rozhranie

Užívateľské rozhranie aplikácie bolo tak ako samotný dátový model vytvárané v softvérovom komponente QlikView desktop, pričom pre lepšiu vizualizáciu údajov a sprehľadnenie jednotlivých okien programu boli pre pozadie aplikácie použité obrázky vytvorené v programe Adobe Photoshop. Keďže QlikView umožňuje rozdelenie takzvanej front-end časti programu na niekoľko blokov, respektíve zošitov, boli na základe vybranej meteorologickej veličiny a logickej štruktúry aplikácie vytvorené nasledujúce okna programu:

- Úvod
- Aktuálne dáta
- Teplota
- Vietor
- Zrážky
- Vlhkosť
- Štatistika

Jednotlivé okna aplikácie zobrazujú príslušné meteorologické údaje, pričom umožnené je tiež zobrazenie aktuálneho stavu počasia v mieste meteorologickej stanice (zošit aktuálne dáta) a to v prípade načítania aktuálnych dát z webových súborov umiestnených na stránkach meteorologickej stanice. Pre toto načítanie je však potrebné v prípade dostupného internetového pripojenia nastaviť načítavanie z webu v úvodnom okne programu. Jednotlivé ukážky vytvorenej aplikácie môžete vidieť na nasledujúcich obrázkoch.



## Analyza meteorologických dát - Písek

Táto aplikácia je zameraná na analýzu meteorologických dát meteorologickej stanice v meste Písek. Meteorologická stanica, ktorej dáta sú analyzované funguje v rámci strednej poľnohospodárskej školy v meste Písek a založená bola v auguste roku 2010 v rámci grantového programu: „Zavádzanie nových technológií do stredných a vyšších odborných škôl“. Po schválení grantu boli objednané jednotlivé prístroje, meteorologická stanica a príslušenstvo potrebné pre samotnú prevádzku, čím sa mohlo prísť k inštalácii. Od septembra 2010 stanica bežala v testovacom režime vnútri budovy a v mesiaci október boli jednotlivé senzory stanice presunuté do exteriéru, kde boli vyhľadane optimálne umiestnenia z hľadiska merania hodnôt, ako aj bezproblémového prenosu dát do radiacej jednotky. Analyzovať dáta je tak možné od mesiaca november 2010 až po úplne aktuálne dáta v prípade načítania dát z webu.



Obr. 2 Úvodne okno aplikácie

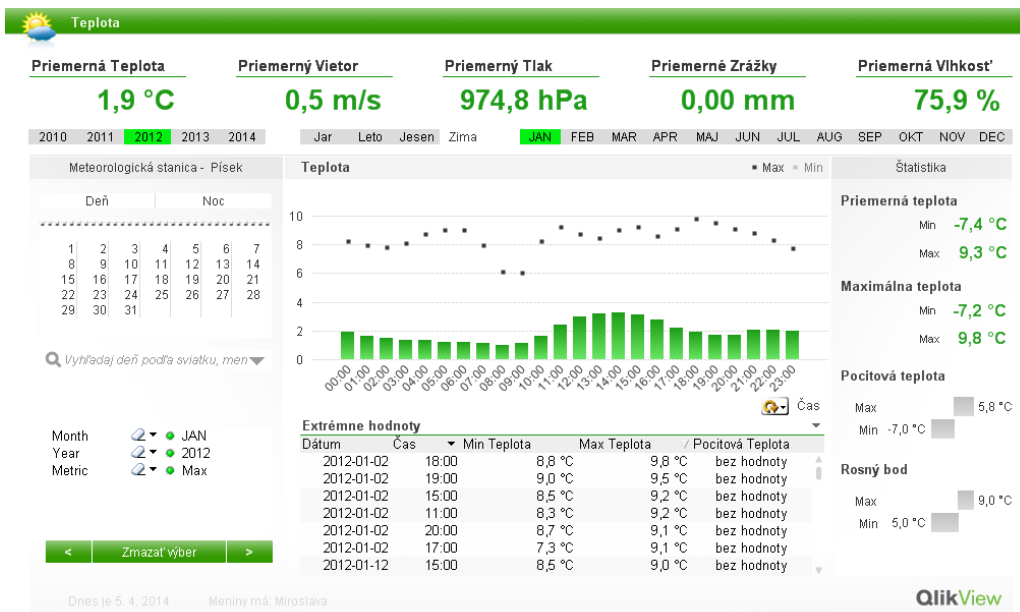
**Aktuálny stav počasia**

<b>Aktuálna Teplota</b> 16,6 °C Priemer 10,2 °C	<b>Aktuálny Vietor</b> 0,1 m/s Priemer 0,6 m/s	<b>Aktuálny Tlak</b> 968,0 hPa Priemer 970,8 hPa	<b>Zrážky za 24 hod.</b> 0,0 mm Priemer 0,0 mm	<b>Aktuálna Vlhkosť</b> 55,0 % Priemer 63,8 %
---	--	--	--	---

Meteorologická stanica - Písek	História	Aktuálne	Ukazovateľ	Aktuálne	História
Dátum a čas posledného merania 5. 4. 2014 13:00	<b>Teplota</b> Max 17,2 °C Min 1,7 °C	-20 °C 0 °C 20 °C 40 °C	<b>Pocitová teplota</b>	- °C	Min - °C Max - °C
Lokalizácia 49° 18' 40" N 49° 18' 40" E Výška: 380 m.n.m.	<b>Vietor</b> Max 2,4 m/s Min 0,0 m/s	0,1 m/s	<b>Rosný bod</b>	8,0 °C	Min -1,0 °C Max 9,0 °C
Prejdí na analýzy	<b>Tlak</b> Max 983,0 hPa Min 982,0 hPa	miestry 9680 hPa prepočet na hl. mora 9675 hPa	<b>Náraz vetra</b>	0,1 m/s	Min 0,0 m/s Max 3,3 m/s
			<b>Zrážky za 24 hod.</b>	0,0 mm	Min 0,0 mm Max 0,1 mm
			<b>Zrážky za 1 hod.</b>	0,0 mm	Min 0,0 mm Max 0,0 mm
			<b>Vlhkosť</b>	55,0 %	Min 45,0 % Max 81,0 %

Dnes je 5. 4. 2014 Meniny má Miroslava **OlikView**

Obr. 3 Okno Aktuálne dáta



Obr. 4 Okno teplota

## V. ZÁVER

Navrhnuté a implementované riešenie informačného systému pre potreby analýzy dát v nástroji QlikView je demonštráciou možnosti a prínosov tejto data discovery platformy, ktorej výhody vidím najmä v jednoduchosti integrácie dát z viacerých zdrojov, asociatívny in-memory dátový model, flexibilitu úprav a priateľivé užívateľské rozhranie.

## POĎAKOVANIE

Táto práca bola podporená grantom KEGA - 021TUKE-4/2012(70%) a grantom VEGA - 1/0286/11 (30%).

## REFERENCIE

- [1] HAJASTEK, P: *Analýza meteorologických dát s využitím nástroja QlikView*. Diplomová práca. Košice: KKUI FEI TUKE, 2014. 83 s.
- [2] QlikTech: QlikTech Delivers Another Quarter of Strong Growth [Online]. [cit. 2013-08-04]. Dostupné na internete: <<http://www.qlikview.com/us/company/press-room/press-releases/2013/en/0725-qliktech-announces-second-quarter-2013-financial-results>>.
- [3] HARMSSEN, Barry – GARCIA, Miguel: *QlikView for developers 11*. Birmingham – Mumbai: Packt Publishing, 2012. 534 s. ISBN 978-1-84968-606-8

# Analýza postojov v neformálnych krátkych textoch

<sup>1</sup>Miroslav KRAVEC, <sup>1</sup>Kristína MACHOVÁ

<sup>1</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita Košice, Slovenská republika

miroslav.kravec@student.tuke.sk, kristina.machova@tuke.sk

**Abstrakt** — Rozvoj informačných a komunikačných technológií v posledných rokoch má za následok enormný nárast množstva informácií, ktoré denne vznikajú vo forme elektronických dokumentov. Pre človeka je náročné takého údaje spracovať. Preto sa v súčasnosti v tomto smere stávajú čoraz častejšie stredobodom pozornosti systémy zamerané na dolovanie znalostí z textov. Jedným z elementov takýchto systémov sú aj oblasti dolovania znalostí z názorov a analýza sentimentu. Proces získavania a extrahovania znalostí z krátkych neformálnych textov vo forme názorov, resp. postojov sa skladá z viacerých krokov. Prvým z nich je získanie a príprava dát na ďalšiu analýzu. Konkrétne sa jedná o príspevky vo forme komentárov, ktoré sa viažu k filmom. Na základe tejto domény boli navrhnuté a vytvorené kategórie jednotlivých druhov postojov. V rámci klasifikácie textu bol využitý slovníkový prístup. Z toho dôvodu potreba vytvoriť dostatočne veľký klasifikačný slovník slov jednotlivých druhov postojov predstavuje jednu z najdôležitejších a najnáročnejších častí práce.

**Kľúčové slová** — postoj, názor, analýza sentimentu, dolovanie znalostí z textov.

## I. ÚVOD

Hlavným prínosom tejto práce je návrh a implementácia prístupu, ktorý v procese rozpoznávania postojov v príspevkoch využíva klasifikačný slovník. Systém extrakcie autoritatívnych postojov zo zdrojových textov je semi-automatizovaný proces. Dôkazom sú postupy a spôsoby získavania dát, tvorby trénovacej množiny ako aj analýza tejto množiny a interpretácia výsledkov.

Z vyššie uvedeného vyplýva, že prvým krokom v procese budovania takéhoto systému je získanie dát. Prvá časť je venovaná opisu získaných dát, ktorými sú príspevky vo forme komentárov. Na to nadväzuje návrh a tvorba jednotlivých kategórií postojov a klasifikačného slovníka. Programové riešenie tohto prístupu predstavuje nielen využívanie potenciálu vytvoreného slovníka, ale aj spracovanie užívateľom zadaného textu a ilustrácia výsledkov vo forme autoritatívneho postoja v rámci tohto textu. Na vyhodnotenie úspešnosti nasadeného systému nad vybranou kolekciou príspevkov boli použité predovšetkým presnosť, návratnosť a F1 skóre.

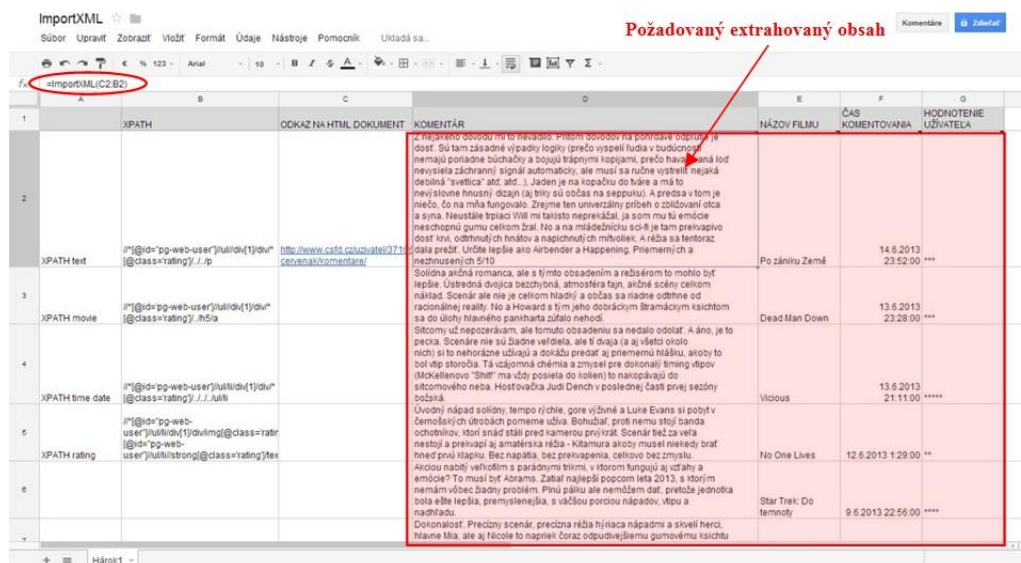
## II. DÁTA

Dátovou základňou v rámci práce zameranej na analýzu postojov sú príspevky k filmom. Konkrétne reálne dáta tvoria komentáre k filmom získané z najznámejšej česko-slovenskej filmovej databázy csfd.cz. Tieto textové zdroje sú tvorené neformálne písanými príspevkami v slovenčine.

Ak chceme vykonať nad nejakým súborom dát príslušnú analýzu, je potrebné tieto dáta v prvom rade získať. Naším cieľom je teda získať textové komentáre k filmom, resp. spracovať zdrojový kód písaný v značkovacom html jazyku do nami žiadanej podoby. Na spracovanie html dokumentov a extrahovanie komentárov bolo použité prostredie "Google Docs Spreadsheet" a to konkrétne funkcia "ImportXML" v tvare `ImportXML("odkaz na html dokument";"XPath príkaz")`. Je to efektívna funkcia vyvinutá spoločnosťou Google, Inc., ktorá na extrahovanie požadovaného obsahu používajú príkazy XPath, ktorým vykonávame navigáciu medzi elementmi a atribútmi v HTML (XML) dokumente.

Na obrázku (Obr.1) môžeme vidieť príklad extrahovania požadovaného obsahu z html dokumentu prostredníctvom funkcie ImportXML. Vidíme, že bunka C2 obsahuje odkaz na nami

zadaný html dokument, stĺpce B2 až B5 sú jednotlivé príkazy xpath-u, ktorým načítame požadovaný obsah.



Obr.1 Príklad obsahu extrahovaného z cieľového html zdroja

### III. TVORBA SLOVNÍKA

Ďalšou významnou súčasťou práce zameranej na analýzu postojov je návrh a budovanie dostatočne rozsiahleho klasifikačného slovníka emocionálnych slov. Existuje niekoľko prístupov pre akvizíciu lexikónu:

**1. Manuálny prístup** - ručne budovaný slovník. Nevýhodou je fakt, že každá doména vyžaduje vlastný špecifický slovník, a preto vývoj a tvorba takéhoto slovníka je priveľmi prácná.

**2. Prístup využívajúci iné slovníky** - ide o spôsob tvorby slovníka využitím zdrojov, ako napríklad WordNet alebo SentiWordNet. Najčastejšie je používaný WordNet, resp. jeho synonymá a antonymá, na základe ktorých je rozšírená počiatočná relatívne malá množina slov. Nevýhodou tohto prístupu je, že je doménovo nezávislý, čiže nezachytáva špecifické osobitosti konkrétnej domény.

**3. Prístup využívajúci korpusedokumentov** - sada slov je budovaná a rozširovaná z dostatočne veľkého korpusedokumentov z danej domény. Odstraňuje nevýhodu v podobe doménovej nezávislosti známej z predošlého prístupu.

V rámci budovania systému analýzy postojov bol použitý prvý prístup. Súčasťou budovania slovníka bolo tiež skúmanie najčastejšie používaných slov v rámci použitej domény ako aj využitie už existujúcich slovníkov emočných slov podľa Ekmanovej teórie rozdelenia emócií, ktoré boli preložené z anglického jazyka do slovenčiny. Slovník pozostáva zo slov s istým emočným nábojom (musia vyjadrovať špecifický postoj), ktorých základ bol manuálne transformovaný na jeho všetky možné tvary - využitie tvaroslovnia.

Každé slovo je zaradené do príslušnej triedy postoja. Návrh jednotlivých kategórií postojov sa odvíjal od použitej dátovej základne analýzy postojov - komentárov k filmom. Tabuľka 1 reprezentuje konkrétne druhy postojov.

Tabuľka 1  
 Kategórie postojov

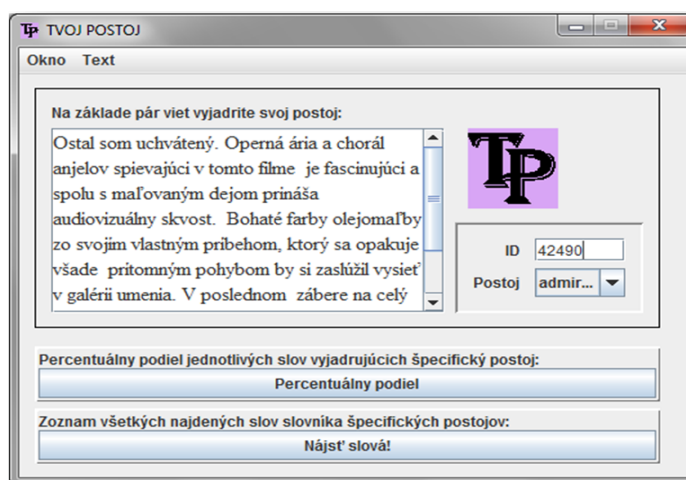
Skratka	Kategória postoja
ADMR	Obdiv - <b>admiration</b>
INTR	Záujem - <b>interest</b>
ACCT	Akceptácia - <b>acceptance</b>
BORD	Unudenosť - <b>boredom</b>
DISG	Znechutenie - <b>disgust</b>
LOAT	Hnus, odpor - <b>loathing</b>

#### IV. IMPLEMENTÁCIA SYSTÉMU

##### A. Opis programového riešenia

Ako vhodné riešenie pre implementáciu systému rozpoznávania postojov v krátkych neformálnych textoch bola zvolená metóda založená na slovníkovom prístupe. Preto aj aplikácia vytvorená v programovacom jazyku JAVA (viď Obr.2) bola vyvinutá za účelom čo najefektívnejšieho využitia potenciálu vytvoreného slovníka postojov. Hlavnou funkciou programu je spracovanie a analýza používateľom zadaného textu v podobe výpočtu percentuálneho vyjadrenia zastúpenia slov navrhnutých druhov postojov (Obr.3).

Dôležitou súčasťou fungovania programu je segmentácia upraveného vstupného textu na samostatné lingvistické jednotky (LJ) - nazývané aj tokeny, ktoré sú pochopiteľne nevyhnutnou súčasťou ďalších procesov zameraných na analýzu postojov. Najprv sú výstupom upraveného textu tokeny ako samostatné slová vzniknuté rozdelením tohto textu na základe medzier. Jednou z funkcií programu je aj možnosť obohatenia množiny slov určených na analýzu o tzv. n-gramy. Konkrétne ide o tvorbu až do úrovne trigramov ( $n = 3$ ).



Obr.2 Ukážka používateľského rozhrania aplikácie

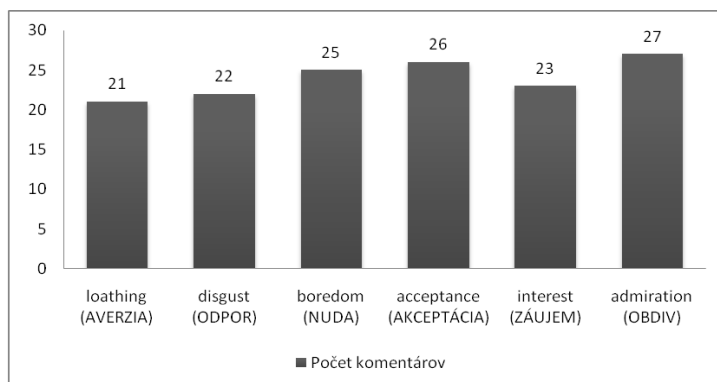
Postoj	Percentuálny podiel
loathing (AVERZIA)	0.0
disgust (ODPOR)	0.0
boredom (NUDA)	0.0
acceptance (AKCEPTÁ...)	0.0
interest (ZÁUJEM)	0.0
admiration (OBDIV)	100.0

ID: 42490  
Postoj experta: admiration

Obr.3 Výstup v podobe percentuálneho zastúpenia jednotlivých druhov postojov

##### B. Tvorba trérovacej množiny

Pri výbere prvkov trérovacej množiny zohrával dôležitú úlohu fakt, aby daný obsah príspevkov vo forme komentárov čo najlepšie zastupoval príslušnú triedu postoja. Podstatné bolo, aby dané textové zdroje obsahovali primeraný počet textového vyjadrenia postojov vo forme slov alebo kratších slovných spojení. Výber bol realizovaný z dostatočne veľkej databázy komentárov (okolo 64 000). Keďže ide predovšetkým o krátke neformálne písané texty, v niektorých prípadoch, ako napríklad "Tak toto bola totalna hardcore nuda. Prve skutočne sklamanie od Jarmusha. No ale každý musí nejako začať." systémom nájdené slová "nuda" a "sklamanie" reprezentujú kvantitatívne malé, avšak primerané a jednoznačné zastúpenie konkrétneho postoja "boredom". Prehľad rozloženia prvkov trérovacej množiny naprieč jednotlivými druhmi postojov znázorňuje graf (Obr.4).



Obr.4 Histogram rozloženia prvkov tréningovej množiny naprieč jednotlivými druhmi postojov

## V. VYHODNOTENIE SYSTÉMU

### A. Možnosti hodnotenia efektívnosti klasifikácie

Meranie efektívnosti klasifikácie sa vykonáva na základe vyčíslenia presnosti  $\pi_j$  (precision) a návratnosti  $\rho_j$  (recall) z kontingenčnej tabuľky (Tabuľka 2):

Tabuľka 2  
 Kontingenčná tabuľka pre kategóriu  $c_j$

	$\Phi(d_i, c_j) = \text{true}$	$\Phi(d_i, c_j) = \text{false}$
$\Phi(d_i, c_j) = \text{true}$	$TP_j$	$FP_j$
$\Phi(d_i, c_j) = \text{false}$	$FN_j$	$TN_j$

$$\pi_j = \frac{TP_j}{TP_j + FP_j}, \rho_j = \frac{TP_j}{TP_j + FN_j} \quad (1)$$

kde  $TP_j$  a  $FP_j$ , resp.  $TN_j$  a  $FN_j$  sú správne a nesprávne klasifikované pozitívne, resp. negatívne príklady triedy  $c_j$ . Presnosť je možné definovať ako pomer počtu správnych výsledkov pozitívnej klasifikácie triedy  $j$  s celkovým počtom výsledkov pozitívnej klasifikácie. Návratnosť je vyjadrená pomerom počtu správnych pozitívnych výsledkov klasifikácie a počtu skutočných pozitívnych príkladov. Kombináciou týchto ukazovateľov miery efektívnosti klasifikátora dosiahneme jeden v podobe harmonického priemeru, tzv. F1 skóre:

$$F_1 = \frac{2 \times \pi \times \rho}{\pi + \rho} \quad (2)$$

Táto funkcia kombinuje presnosť a návratnosť s rovnakou váhou. Kombinovať hodnoty presnosti a návratnosti môžeme do miery  $F_\beta$  na základe vzťahu:

$$F_\beta = \beta^2 + 1 \times \frac{\pi \rho}{\beta^2 \pi + \rho} \quad (3)$$

pričom parameter  $\beta$  vyjadruje mieru vplyvu  $\pi$  a  $\rho$ .

### B. Vyhodnotenie úspešnosti systému analýzy postojov

Bol realizovaný výpočet ukazovateľov efektívnosti nasadeného systému a to presnosti, návratnosti a skóre F1 pre každú triedu (kategóriu) zvlášť. Najvyššia dosiahnutá presnosť klasifikácie bola dosiahnutá v rámci triedy "admiration" - 77,8 %. Pri ukazovateli F1 skóre, kombináciou hodnôt presnosti a návratnosti bol dosiahnutý najlepší výsledok taktiež v rámci triedy "admiration" - 79,2 %. Naopak, najmenej úspešná z pohľadu tohto ukazovateľa bola klasifikácia v rámci kategórie "acceptance" - 57,1 %. Najnižšia presnosť klasifikácie bola dosiahnutá v kategórii "interest" - 60,9 %. Tabuľka 3 znázorňuje výsledné priemerné hodnoty sledovaných ukazovateľov.

Tabuľka 3  
Priemerné hodnoty ukazovateľov

	Presnosť	Návratnosť	F1 skóre
AVG (%)	68.06	74.06	70.76

## VI. ZÁVER

Hlavným cieľom vyššie popísanej aplikácie, vytvorenej v programovacom jazyku Java, bolo na základe vytvoreného slovníka určiť kategóriu postoja vyjadreného v rámci užívateľom zadaného príspevku. Tento cieľ bol splnený primerane úspešne s výslednou priemernou presnosťou 68 percent a úspešnosťou 74 percent. Autoritatívny postoj je vyjadrený prostredníctvom percentuálneho podielu jednotlivých slov, ktoré sú súčasťou príspevku.

Proces identifikácie postoja bol aplikovaný nad vybranou kolekciou príspevkov z veľkej databázy komentárov. Bolo potrebné, aby tieto vybrané krátke texty čo najlepšie reprezentovali jednotlivé triedy postojov. Následne bolo vykonávané vyhodnotenie efektívnosti systému analýzy postojov využitím ukazovateľov presnosti, návratnosti, F1-skóre a odhadu chyby.

V budúcnosti by bolo vhodné pracovať na zvýšení efektívnosti aplikácie, predovšetkým rozširovaním použitého slovníka a spresňovaním významu a kategorizácie slov v tomto slovníku do jednotlivých tried postoja. Taktiež by mohlo byť zaujímavé do budúcnosti obohatiť aplikáciu o možnosť semi-automatického fungovania aplikácie, ktorá by v prípade nejasnosti daného slova ponúkla možnosť korekcie používateľom.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## ZOZNAM POUŽITEJ LITERATÚRY

- [1] R. Feldman, J. Sanger, „*Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*“, Cambridge University Press, New York, NY, USA. 2006. ISBN-13 978-0-511-33507-5.
- [2] R. Feldman, „*Techniques and applications for sentiment analysis*“. *Communications of the ACM*, 2013, 56.4: 82-89.
- [3] K. Machova, „*Strojové učenie v systémoch spracovania informácií*“. ELFA s.r.o., 2009, Košice, 85s, ISBN 978-80-8086-130-8.
- [4] K. Machova, P. Bednár, M. Mach, „*Various Approaches to Web Information Processing*“. *Computing and Informatics*, Vol. 26, No. 3, 2007, Bratislava, 301-327, ISSN 1335-9150.

# Vyhľadávanie názorového spamu

Zuzana NEMČIŠINOVÁ, Kristína MACHOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

zuzana.nemcisinova@student.tuke.sk, kristina.machova@tuke.sk

**Abstrakt** — Táto práca sa zaoberá problematikou dolovania v konverzačnom obsahu a vyhľadávaním názorového spamu. Cieľom bolo vytvoriť aplikáciu, ktorá na základe slovníkovej metódy porovnávania slov dokáže identifikovať názorový spam. Aplikácia obsahuje jednoduché používateľské rozhranie a je naprogramovaná v jazyku Java. Súčasťou tejto aplikácie je dostatočne rozsiahly klasifikačný slovník emocionálnych slov, ktorý je základom pre použitie slovníkovej metódy. V závere praktickej časti je vykonané testovanie danej aplikácie a vyhodnotenie úspešnosti navrhnutého riešenia.

**Kľúčové slová** — Webové diskusné skupiny, klasifikačný slovník, konverzačný obsah, názorový spam, slovníková metóda

## I. ÚVOD

Rozvoj modernej spoločnosti spôsobil, že internet a počítače sa stali našou neoddeliteľnou súčasťou. Internet je celosvetová počítačová sieť, ktorá sa stala silným komunikačným nástrojom, ktorý sa neustále šíri po celom svete. Internetová komunikácia priniesla so sebou možnosť v jednom okamihu komunikovať s viacerými ľuďmi, dávať otázky naraz tisíckam používateľom alebo im nechať odkaz. Internet umožňuje ľuďom komunikovať rôznymi spôsobmi.

Z dôvodu neustále sa zvyšujúcej internetovej komunikácie je čoraz viac potrebná detekcia názorového spamu. Navrhnutá metóda vyhľadávania názorového spamu je súčasťou riešenie problému analýzy názorov v konverzačnom obsahu. Príspevok sa venuje spôsobom dolovania v konverzačnom obsahu a zameriava sa na návrh metódy identifikácie názorového spamu, jeho implementovanie, testovanie vytvorenej implementácie a vyhodnotenie aplikácie.

## II. KONVERZAČNÝ OBSAH V INTERNETOVEJ KOMUNIKÁCI

O komunikáciu medzi používateľmi webu, teda v kybernetickom priestore, má záujem čoraz viac ľudí. Najviac sa zvyknú používať rozličné platformy sociálneho webu ako rozličné sociálne siete, rôzne diskusné skupiny, fóra, blogy a podobne. Prostredníctvom týchto webových služieb ľudia medzi sebou komunikujú ale aj sa ovplyvňujú sa. Taktiež môžu vyjadriť svoj názor formou príspevkov, komentárov, recenzií na rôzne témy. [1] Obsah, ktorý bol vytvorený používateľmi v rámci takzvaného konverzačného obsahu zahŕňa v sebe cenné informácie. Na druhej strane dôveryhodnosť konverzačného obsahu nie je nikdy zaručená. Web sám o sebe neobsahuje žiadne služby na kontrolu kvality konverzačného obsahu. A tak sa web stáva vhodným prostredím pre tvorbu okrem iného aj názorového spamu. [2] Väčšina existujúcich antispamových programov sa zameriava iba na detekciu e-mailových spamov. Cieľom prezentovanej práce bolo vytvoriť aplikáciu na vyhľadávanie uvedeného názorového spamu v konverzačnom obsahu.

Čo je to vlastne názorový spam? Sociálne siete umožňujú ľuďom medzi sebou komunikovať a vyjadrovať svoj názor prostredníctvom rôznych komentárov, príspevkov, správ alebo sa pridávať do skupín s ľuďmi, ktorí majú podobné záujmy alebo názory. Práve na sociálnych sieťach sa hromadí obrovské množstvo konverzačného obsahu vytváraného používateľmi. Tento konverzačný obsah okrem cenných informácií obsahuje aj bezcenné príspevky, ktoré neposkytujú žiadny názor a teda sú pre používateľa bezvýznamné. Takto definovaný príspevok môžeme označiť ako názorový spam. [3]



### III. VYHLADÁVANIE NÁZOROVÉHO SPAMU

Na vyhľadávanie názorového spamu bola navrhnutá metóda a aplikovaná, ktorá využíva slovníkovú metódu, prostredníctvom ktorej sa porovnáva vstupný text so slovami zo slovníka. Je to metóda relatívne jednoduchá ale dosahuje pomerne dobré výsledky. K tejto metóde bolo potrebné navrhnuť dostatočne veľký slovník emočných slov, na základe ktorého je možné vykonať detekciu názorového spamu.

#### A. Návrh klasifikačného slovníka

Spolu s viacerými študentmi som sa podieľala na tvorbe slovenského slovníka emotívne ladených slov vyjadrujúcich názor, postoj alebo sentiment. Tento slovník obsahuje podstatné mená, prídavné mená, slovesá, príslovky vo všetkých tvaroch.

Slovník obsahuje 19 403 slov a každému slovu je pridelený stupeň polarizácie názoru. Keďže tento slovník je zameraný na sentiment a obsahuje aj neutrálne slová bolo potrebné upraviť pre potreby detekcie názorového spamu.

Tento slovník bol prejdený manuálne a boli odstránené neutrálne slová (nulový stupeň polarizácie názoru) a tie, ktoré neobsahovali žiadnu emóciu. Taktiež bola vynechaná informácia o type sentimentu (type emócie), ktorá bola nadbytočná pre potreby našej aplikácie, ktorá funguje na jednoduchom princípe porovnávania slov. Po nájdení zhody medzi slovami v príspevku a slovami v slovníku je spracovávaný príspevok vyhodnotený ako užitočný.

#### B. Obsah klasifikačného slovníka

V konečnom dôsledku slovník emocionálnych slov, ktorý slúži na vyhľadávanie názorového spamu obsahuje 31 678 slov. Tento slovník obsahuje vyskloňované podstatné mená, vystupňované prídavné mená, vyčasované slovesá a rôzne druhy prísloviak. Keďže mnoho ľudí pri písaní príspevkov nepoužíva diakritiku bol tento slovník rozšírený o slová bez diakritiky. Čím je slovník väčší, tým je detekcia názorového spamu presnejšia.

#### C. Implementácia riešenia

Navrhnutá slovníková metóda na identifikáciu názorového spamu bola implementovaná v jazyku Java. Tak vznikla aplikácia, ktorá na základe vloženého príspevku a následného porovnávania slov so slovníkom je schopná poskytnúť používateľovi výsledok v podobe informácie, že či daný príspevok je spamový alebo užitočný. Vyhodnotenie príspevku prostredníctvom aplikácie sa skladá z 3 krokov:

1. vloženie a spracovanie textu zadaného používateľom,
2. porovnávanie textu so slovníkom,
3. vyhodnotenie textu.

### IV. TESTOVANIE APLIKÁCIE

Na testovanie aplikácie, vytvorenej v tejto práci boli použité 2 testovacie množiny.

Prvou sú technické príspevky, ktoré sú zamerané na témy ako sú mobily, počítače, notebooky a internet. Tieto príspevky boli čerpané z komentárov, recenzií a diskusných skupín, ktoré je možné nájsť na stránkach: „[www.pretaktovanie.zoznam.sk](http://www.pretaktovanie.zoznam.sk)“, „[www.tech.sme.sk](http://www.tech.sme.sk)“ a „[www.pcforum.sk](http://www.pcforum.sk)“. Táto testovacia množina s bežnými príspevkami obsahuje 5682 slov, ktoré tvoria obsah 130 príspevkov.

Druhou testovacou množinou sú bežné príspevky. Ide o príspevky zaoberajúce sa témou, ako je zdravie, krása, politika a produkty. Tieto príspevky je možné nájsť na týchto stránkach: „[www.zdravie.sk](http://www.zdravie.sk)“, „[www.topky.sk](http://www.topky.sk)“ a „[www.diva.aktuality.sk](http://www.diva.aktuality.sk)“. Táto testovacia množina obsahuje 130 príspevkov, ktoré obsahujú 3771 slov.

Celkovo bolo otestovaných 260 príspevkov. Testovanie bolo založené na porovnávaní príspevkov aplikáciou na vyhľadávanie názorového spamu a expertom (človekom), na základe jeho vlastného uváženia a názoru.

Tab. 1 znázorňuje kontingenčnú tabuľku, v prípade vyhodnotených technických príspevkov. Na základe tejto kontingenčnej tabuľky bola vypočítaná presnosť aplikácie  $\pi=64,71\%$  a návratnosť aplikácie  $\rho=73,33\%$ .

Tab. 2 znázorňuje kontingenčnú tabuľku, v prípade vyhodnotených bežných príspevkov. Na základe tejto kontingenčnej tabuľky bola vypočítaná presnosť aplikácie  $\pi=74,60\%$  a návratnosť  $\rho=79,66\%$ .

Tab. 1  
 Kontigenčná tabuľka - technické príspevky

SPAM: A/N		EXPERT	
		A	N
APLIKÁCIA	A	44	24
	N	16	46

Tab. 2  
 Kontigenčná tabuľka - bežné príspevky

SPAM: A/N		EXPERT	
		A	N
APLIKÁCIA	A	47	16
	N	12	55

Pri porovnaní týchto dvoch testovacích množín je jasné, že aplikácia dosahuje lepšiu presnosť pri vyhodnocovaní bežných príspevkov. Aplikácia v prípade týchto príspevkov dosahuje presnosť 74,60 % a v prípade technických príspevkov je to 64,71 %. To môže byť spôsobené tým, že technické príspevky majú často informatívny charakter a nenesú žiadny názor, aj napriek tomu, že aplikácia ich označila za užitočné. Aplikácia sa mohla pomýliť aj v prípadoch kedy príspevky obsahujú opytovacie vety. Keďže aplikácia je naprogramovaná tak, že ak príspevek obsahuje otáznik, aplikácia automaticky označuje daný príspevek za spamový, bez ohľadu na to aké sú ostatné slová príspevku. Ide väčšinou o príspevky s väčším počtom slov, ktoré môžu obsahovať opytovacie vety, ale často poskytujú aj cenné informácie vyjadrujúce názor k danej téme.

Nízka presnosť môže byť spôsobená aj tým, že niektoré diskusné skupiny sú riadené. To znamená, že používatelia majú prístup ku konverzácii až po zaregistrovaní sa na danej stránke, čím sa zaväzujú dodržiavať stanovené podmienky. V prípade ich porušenia je používateľ z diskusie vylúčený. Niektoré diskusné skupiny umožňujú vymazať príspevky používateľov, ktorí porušili podmienky, prípadne vymazať všetky príspevky používateľa, ktorý opätovne porušoval podmienky. A práve do tejto skupiny spadajú aj príspevky obsahujúce názorový spam. Tieto konverzačné stránky spravujú administrátori, ktorí kontrolujú všetky pridané príspevky. Tie musia spĺňať všetky stanovené podmienky a nesmú odbočovať od témy, inak sú vymazané. [4]

## V. ZÁVER

Aj keď existuje obrovské množstvo antispamových programov, nie je možné odstrániť spam z bežného života úplne. Bohužiaľ spam prenikol do e-mailových správ, mobilných telefónov a najnovšie je ho možné nájsť v rôznych diskusných skupinách, na sociálnych sieťach, v recenziách produktov, v ktorých je možné detekovať spam iba manuálne. Práve takouto problematikou sa zaoberá, táto práca. Výsledkom je aplikácia implementovaná v jazyku Java. Výsledná aplikácia obsahuje jednoduché používateľské rozhranie. Je ľahko ovládateľná a funguje na princípe porovnávania slov.

Táto aplikácia dosahuje pomerne dobré výsledky. Na základe 2 testovacích množín, ktoré boli zamerané na rôzne témy a obsahovali 260 príspevkov bola zistená presnosť aplikácie 74,60 %, v prípade bežných príspevkov. Druhou skupinou boli technické príspevky, ktoré dosiahli presnosť 64,71 %, čo je o 9,89 % nižšia presnosť ako u bežných príspevkov. To môže byť spôsobené tým, že klasifikačný slovník je zameraný iba na užitočné slová vyjadrujúce názor. Zároveň, by bolo vhodné tento klasifikačný slovník navrhnutý pre túto prácu ešte raz manuálne prekontrolovať, respektíve rozšíriť ho o nové slová. Nízka presnosť môže byť spôsobená aj ošetrením otáznika v príspevkoch, taktiež tým, že v súčasnosti už existujú riadené diskusné skupiny.

V budúcnosti je potrebné sa zamerať na zlepšenie presnosti napríklad rozšírením klasifikačného slovníka obsahujúceho emočné slová a vytvorením slovníka najčastejšie sa vyskytujúcich spamových slov. K presnosti by dopomohlo aj priradenie stupňa sentimentu (polarity názoru) k slovám v daných slovníkoch. Taktiež by bolo potrebné ošetriť počet výskytov otáznikov vzhľadom na dĺžku príspevku.

Aj keď táto aplikácia dosahuje nadpolovičnú presnosť, nie je však natoľko presná, aby mohla byť použitá ako hlavný program na detekciu názorového spamu. Je potrebné popracovať na zdokonalení tejto aplikácie. Avšak túto aplikáciu je aj napriek tomu možné použiť ako istú pomoc pri rozhodovaní sa o užitočnosti príspevku.

#### POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### ZOZNAM POUŽITEJ LITERATÚRY

- [1] “Elektronická komunikácia” [online] 2013-07-10. Dostupné na internete: <<http://www.pulib.sk/kis/kacmarikova/1rocnik/I1/I8.pdf>>.
- [2] J. Štofa, P. Michalik, “Conversational content in the context of safety of social networks” [online] 2013-07-10. Dostupné na internete: <<http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6617576&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel7%2F6598373%2F6617553%2F06617576.pdf%3Farnumber%3D6617576>>.
- [3] “Sociálna sieť” [online]. 2014-04-28. Dostupné na internete: <[http://sk.wikipedia.org/wiki/Soci%C3%A1lna\\_sie%C5%A5\\_\(internet\)](http://sk.wikipedia.org/wiki/Soci%C3%A1lna_sie%C5%A5_(internet))>.
- [4] M. Senedek, “Dynamická zmena autority aktérov sociálneho webu” Diplomová práca. Košice, 2013. 16-17 s.

# Blocks of Cognitive Radio System

<sup>1</sup>Štefan MASTILÁK, <sup>2</sup>Lukáš SENDREI, <sup>3</sup>Stanislav MARCHEVSKÝ

<sup>1,2,3</sup> Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>stefan.mastilak@student.tuke.sk, <sup>2</sup>lukas.sendrei@tuke.sk,  
<sup>3</sup>stanislav.marchevsky@tuke.sk

**Abstract** — Cognitive radio is an emerging and fascinating technology for the opportunistic use of under-utilized spectrum that promises to change the future technological trends forever if employed properly. This technique allows for usage of licensed frequency bands by unlicensed users. However, these unlicensed (cognitive) users need to monitor the spectrum continuously to avoid possible interference with the licensed (primary) users. This work explains the background of cognitive radio technology and focuses on the main blocks of a cognitive radio.

**Keywords** — cognitive radio, spectrum sensing, spectrum management, spectrum sharing

## I. INTRODUCTION

Today's wireless networks are regulated by a fixed spectrum assignment policy, i.e., the spectrum is regulated by national regulatory organization, such as Federal Communications Commission (FCC). The FCC assigns spectrum to licensed holders, also known as primary users (PUs), or services on a long term basis for large geographical regions. However, a large portion of the assigned spectrum remains underutilized. An example of this fact is depicted in Fig. 1.

The inefficient utilization of the limited spectrum necessitates the development of dynamic spectrum access (DSA) technologies, where users who have no spectrum licenses, also known as secondary users (SUs), are allowed to use the temporarily unused licensed spectrum. In recent years, the FCC has been considering more flexible and comprehensive uses of the available spectrum, through the use of cognitive radio (CR) technology. CR technology is envisaged to solve the problems in wireless networks resulting from the limited available spectrum and the inefficiency in the spectrum usage by exploiting the existing wireless spectrum opportunistically [1]-[4].

The rest of the paper is organized as follows. The second section gives a short overview of the CR terminology and key characteristics. The third section describes the functions of the three main blocks of CR and the interactions between them. The concept of spectrum sensing in general and, as well, some sensing techniques are depicted in the fourth section. Spectrum management and spectrum sharing are analyzed in the fifth and sixth section, respectively. Finally, the last section gives the conclusions.

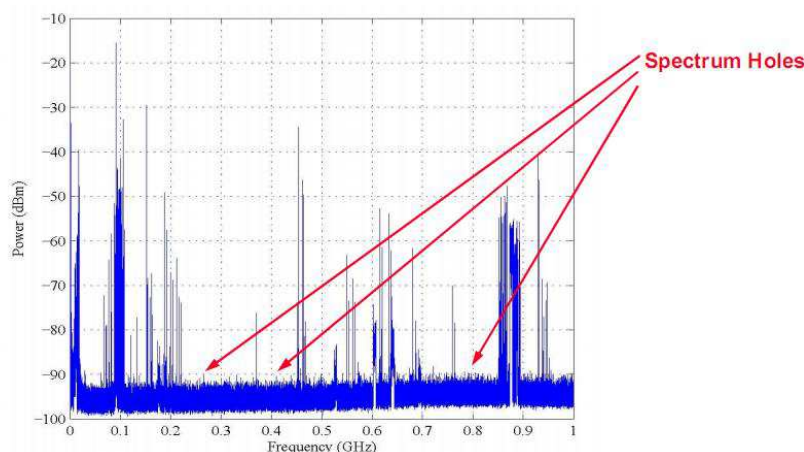


Fig. 1 Underutilized spectrum (spectrum holes or white spaces).

## II. COGNITIVE RADIO

CR is designed to identify and scavenge the spectrum holes in the licensed spectrum bands as shown in Fig. 2. A spectrum hole is defined as a licensed spectrum band that can be used by CR users without interfering the primary or licensed users. Generally spectrum holes can be broadly divided into two categories: temporal spectrum holes and spatial spectrum holes. A temporal spectrum hole means that there is no primary transmission over the spectrum band of interest during the time of sensing (over a reasonable period); hence, this band can be utilized by CR users in the current time slot. A spatial spectrum hole exists when the spectrum band of interest is occupied by the primary transmission only in a restricted area; hence, this band can be utilized by CR users well outside this area [6].

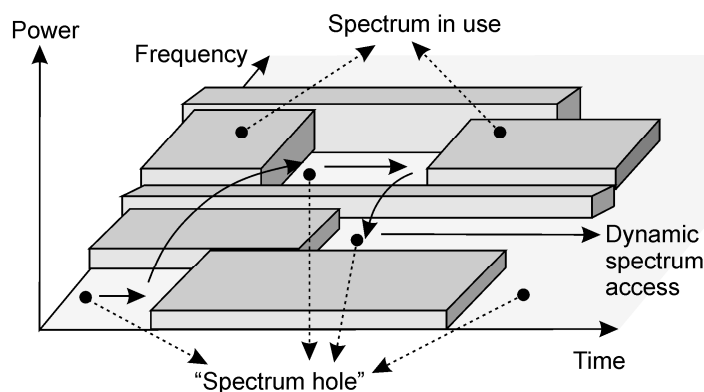


Fig. 2 Dynamic spectrum access and spectrum hole concept [2].

The two important characteristics of cognitive radio are:

- a) Cognitive capability - Refers to the ability of the cognitive radio technology to capture or sense the information from its radio environment.
- b) Reconfigurability - Enables the cognitive radio to be programmed dynamically according to the radio environment [2].

The CR is considered to be the key enabling technology that enables next generation (XG) communication networks, also known as DSA networks, to utilize the spectrum more efficiently in an opportunistic fashion without interfering with PUs [1].

DSA technologies allow CR to operate in the best available channel. More specifically, the CR technology will enable the users to:

1. Determine which portions of the spectrum is available and detect the presence of licensed users when a user operates in a licensed band (spectrum sensing)
2. Select the best available channel (spectrum decision)
3. Coordinate access to this channel with other users (spectrum sharing)
4. Vacate the channel when a licensed user is detected (spectrum mobility) [1].

Once a CR supports the capability to select the best available channel, the next challenge is to make the network protocols adaptive to the available spectrum. Hence above four spectrum management functions are required in an XG communication networks to support this adaptivity. These functions of XG networks enable spectrum-aware communication protocols. Considering the more flexible and comprehensive use of the spectrum resources, especially when SUs coexist with PUs, traditional spectrum allocation schemes and spectrum access protocols may no longer be applicable. So, new spectrum management approaches need to be developed to solve new challenges in research related to CR, specifically, in spectrum sensing and dynamic spectrum sharing [1].

## III. BLOCKS OF COGNITIVE RADIO

Cognitive radio signifies a radio that employs model-based reasoning to achieve a specified level of competence in radio-related domains [5]. Cognitive radio architecture is illustrated in Fig. 3 which consists of three main blocks. Spectrum sensor network provides sensing of signal spectrum and chooses the best channel for secondary user if there is no primary licensed user at current time. Best available channel is selected by decision algorithm. Cognitive Radio network communicate with Licensed network and vacate the channel when a licensed user is detected.

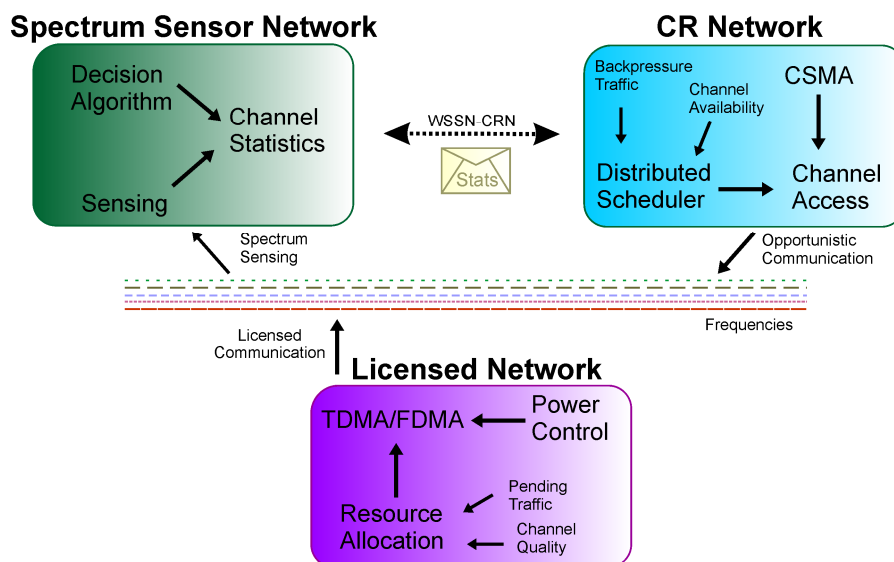


Fig. 3 Block scheme of CR

#### IV. SPECTRUM SENSING TECHNIQUES

As secondary users, CR operators are allowed to utilize a licensed band only when they do not cause interference to the primary users. Spectrum sensing aims at monitoring the usage and characteristics of the covered spectral band(s) and are thus required by CR users both before and during the use of licensed spectrum bands [6].

There are various spectrum sensing techniques which are employed for spectrum sensing:

- 1) Energy Detection.
- 2) Matched Filter Detection.
- 3) Primary signal Detection.
- 4) Cyclostationary Detection.
- 5) Cooperative Detection.

All these methods are good and each has its own advantages and disadvantages. When the primary user signal is unknown or the receiver cannot gather sufficient information about the primary user signal, the energy detection method is used. This method is the simplest and has low computational and implementation complexities and optimal under no knowledge of primary signal.

An energy detector (ED) simply treats the primary signal as noise and decides on the presence or absence of the primary signal based on the energy of the observed signal. Since it does not need any a priori knowledge of the primary signal, the ED is robust to the variation of the primary signal. Moreover, the ED does not involve complicated signal processing and has low complexity. In practice, energy detection is especially suitable for wide-band spectrum sensing. In this case, the simultaneous sensing of a number of subbands can be realized by simply scanning the power spectral density (PSD) of the received wide-band signal. In practice, it is advisable to complete wide-band spectrum sensing via two stages. In the first stage, low-complexity energy detection is applied to search for possible idle subbands; in the second stage, more advanced spectrum sensing techniques with a higher detection sensitivity and thereby higher complexity, such as cyclostationary detection, are applied to the subband candidates to determine whether they are actually available for secondary usage [6].

#### V. SPECTRUM MANAGEMENT

Spectrum management assigns the available spectrum holes among the cognitive radios and is one of the main challenges in cognitive radio [3]. In cognitive radio networks, the unused spectrum bands will be spread over wide frequency range. These unused spectrum bands detected through spectrum sensing show different characteristics according to the time varying radio environment and the spectrum band information such as the operating frequency and bandwidth. So, now the cognitive radio networks should select the available unused spectrum bands and assign them to the secondary users. The selection of unused spectrum bands by cognitive radio network is known as spectrum management [2].

As shown in Fig. 4, the steps of the cognitive cycle consist of four spectrum management functions: spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility. The followings are the main features of spectrum management functions [1].

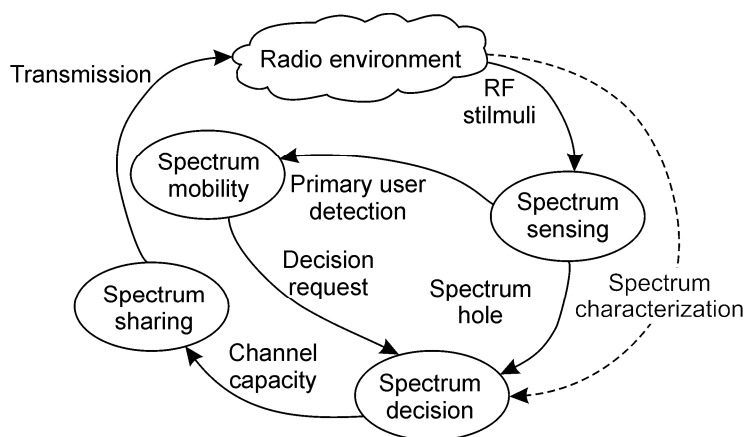


Fig. 3 Block scheme of CR

## VI. SPECTRUM SHARING

Since there may be multiple SUs trying to access the spectrum, their transmissions should be coordinated to prevent collisions in overlapping portions of the spectrum. Spectrum sharing provides the capability to share the spectrum resource opportunistically with multiple CR users which includes resource allocation to avoid interference caused to the primary network. For this, game theoretical approaches have also been used to analyze the behavior of selfish SUs. Furthermore, this function necessitates a CR MAC protocol, which facilitates the sensing control to distribute the sensing task among the coordinating nodes as well as spectrum access to determine the timing for transmission [1].

## VII. SPECTRUM MOBILITY

If a PU is detected in the specific portion of the spectrum in use, SUs should vacate the spectrum immediately and continue their communications in another vacant portion of the spectrum. For this, either a new spectrum must be chosen or the affected link may be circumvented entirely. Thus, spectrum mobility necessitates a spectrum handoff scheme to detect the link failure and to switch the current transmission to a new route or a new spectrum band with minimum quality degradation. This requires collaborating, and routing protocols. Furthermore, this functionality needs a connection management scheme to sustain the performance of upper layer protocols by mitigating the influence of spectrum switching [1].

## VIII. CONCLUSION

We described the background of CR technology, main blocks of CR network and basic spectrum sensing methods in this paper. Dynamic allocation of white spaces in spectrum is very important when designing CR. Efficiency of spectrum allocation determines the throughput of network as well as the total spectrum utilization. Considering the continually increasing amount of data traffic in wireless networks, there is a huge need for better spectrum policies. Therefore, CR wireless networks can be an attractive solution to the problem of spectrum underutilization.

## ACKNOWLEDGMENT

This work was supported by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the contract no. 1/0766/14.

## REFERENCES

- [1] K. Bosung, K. Sehoon, M.T. Nguyen, K. Jong-yon, R. Byeong-hee, Ch. Jungwook, B. P. Soo: "Modeling and Simulation for Cognitive Radio in OPNET". Graduate School of Information and Communication, Ajou University, Suwon-si, Gyeonggi-do, Korea, pp. 1-4.
- [2] A. Bansal, R. Mahajan: "Building a Cognitive Radio System Using Matlab", International Journal of Electronics and Computer Science Engineering, ISSN 2277-1956/V1N3-1555-1560, pp. 1-3.
- [3] F. Khozeimeh and S. Haykin: "Self-organizing dynamic spectrum management for cognitive radio networks," 8th Annual Communication Networks and Services Research Conference, pp. 1-8, 2010.
- [4] B. Wang and K. J. R. Liu: "Advances in Cognitive Radio Networks: A Survey," IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 1, February 2011, pp. 5-23.
- [5] J. Mitola, III, "Cognitive Radio for Flexible Multimedia Communications," Mobile Multimedia Communications, 1999(MoMuC'99) 1999 IEEE International Workshop on, pp. 3-10, 1999.
- [6] B. Jun Ma, G. Ye Li, B. Hwang (Fred) Juang, Fellow IEEE: "Signal Processing in Cognitive Radio", IEEE 97, No. 5, May 2009

# Návrh a implementácia paralelných programov s využitím JCSP

<sup>1</sup>Martin ŠOLC, <sup>2</sup>Slavomír ŠIMONĀK

<sup>1</sup> Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>2</sup> Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>martin.solc@student.tuke.sk, <sup>2</sup>slavomir.simonak@tuke.sk

**Abstrakt** — Tento článok je venovaný návrhu a implementácii paralelných programov s využitím JCSP. Vybraný paralelný systém je najskôr implementovaný, na základe formálnej definície algebrou CSP, s využitím knižnice JCSP. Následne je implementovaný bez použitia knižnice JCSP. Dva vytvorené varianty implementácie sú následne porovnávané z rôznych hľadísk.

**Kľúčové slová** — procesná algebra, verifikácia, CSP, JCSP, ABP

## I. ÚVOD

Pri paralelnom programovaní a súčasnej vzájomnej interakcii vlákien sa systém môže dostať do niektorého z nežiaducich stavov, napríklad:

- Vzájomné čakanie (*deadlock*) – skupina vlákien sa dostane do stavu, kde každé vlákno čaká na iné, aby mohlo pokračovať v činnosti.
- Uviaznutie (*livelock*) – skupina vlákien sa dostane do nekonečného cyklu vzájomných interakcií bez reakcie na vonkajšie podnety.
- Vyhladovanie (*starvation*) – nekonečné odloženie alebo permanentné blokovanie jedného alebo viacerých spustených vlákien.
- Súťaženie (*race hazard*) – dve alebo viac vlákien sa snažia získať prístup k rovnakým zdieľaným údajom a aspoň jeden z prístupov má byť operácia zapisovania.

Mnohé z týchto hazardov je ťažko odhaliť. Pri testovaní systému sa nemusia vôbec vyskytnúť, no môžu sa prejaviť neskôr pri používaní systému. Aj po objavení problému v systéme, môže byť veľmi zložitá ho identifikovať a vôbec lokalizovať.

Spomenuté problémy je možné odhaliť a predísť im využitím formálnych metód pri návrhu a implementácii paralelných programov. Systém sa formálne definuje a následne verifikuje, čo umožňuje vylúčiť prechod systému do stavov, ktoré obmedzia jeho funkčnosť. Systém definovaný procesnou algebrou CSP je možné implementovať pomocou knižnice JCSP. Článok prináša porovnanie návrhu a implementácie paralelného systému s využitím knižnice JCSP a bez jej využitia.

## II. FORMÁLNE METÓDY

Formálne metódy sú techniky využívajúce striktné špecifikované matematické modely na navrhovanie softvérových a hardvérových systémov [1]. Na zaistenie spoľahlivosti a korektného správania sa systému sa využívajú pri testovaní matematické dôkazy. Jedná sa o dôkladnejšiu metódu overenia systému ako je len empirické testovanie. Klasické testy potvrdia len situácie, v ktorých systém pracuje správne [1]. Nemôžu však určiť správanie sa systému v situáciách nezahrnutých v testoch. Formálna verifikácia neodstraňuje zlé predpoklady v návrhu.

### A. Procesná algebra

Procesná algebra je teda štúdiom správania paralelných alebo distribuovaných systémov pomocou algebrických prostriedkov. Vďaka nej možno popísať takéto systémy. Elementom procesnej algebry je proces, nad ktorým pomocou axióm možno prevádzať rôzne výpočty. Pomocou jej prostriedkov možno opísať paralelnú kompozíciu, takisto možno popísať aj



kompozíciu alternatív, teda voľbu, a sekvenčnú kompozíciu [2]. Pomocou procesnej algebry je tak možné popísať aj celé systémy. Medzi najznámejšie procesné algebry patria:

- CCS (Calculus of Communicating Systems),
- ACP (Algebra of Communicating Processes),
- CSP (Communicating Sequential Processes).

#### CSP

CSP (Communicating Sequential Processes) je matematická teória na špecifikáciu a verifikáciu komplexných vzorov správania, vyplývajúceho zo vzájomného pôsobenia medzi súbežnými, paralelnými objektmi [3]. Túto teóriu vyvinul Tony Hoare [4]. CSP má kompozičnú sémantiku, ktorá značne zjednodušuje návrh a tvorbu takýchto systémov. CSP model bol zavedený aj do programovacieho jazyka Java. Bol implementovaný ako knižnica balíkov, ktorá umožňuje, aby boli viacvláknové systémy navrhované a implementované výhradne z hľadiska CSP synchronizačných primitív (ako sú kanály, udalosti, atď.) a konštruktorov (paralelizmus, voľba, atď.) [3].

### III. JCSP

JCSP (Communicating Sequential Processes for Java) poskytuje kompletnú knižnicu na vytvorenie komplexnej funkcionality pomocou siete pozostávajúcej z komunikujúcich procesov [5]. JCSP je v súlade s CSP modelom komunikujúcich systémov, ktorého teória, nástroje a praktické skúsenosti boli pretavené do podpory viacvláknových aplikácií v Jave.

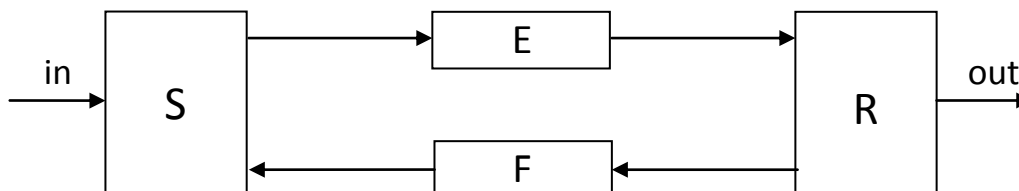
Komunikácia a synchronizácia medzi procesmi prebieha výhradne prostredníctvom malej množiny synchronizačných primitív, ktoré ponúka CSP, alebo prostredníctvom iných starostlivo definovaných spôsobov prístupu k zdieľaným pasívnym objektom [6]. Procesy nevolajú navzájom svoje metódy. Každý proces riadi svoj vlastný stav a väzbu s okolitým prostredím má len prostredníctvom dostupných primitív. Procesy môžu byť skombinované a spustené ako postupnosť, teda sekvenčná kompozícia dvoch procesov, alebo môžu byť spustené paralelne, ako paralelná kompozícia procesov.

JCSP je akousi alternatívou k modelu monitora vlákien v jazyku Java. Model monitora vlákien, ponúkaný v programovacom jazyku Java, je síce jednoduchý na porozumenie, ale je ťažké ho bezpečne použiť v komplexnejšom systéme. Jedným z problémov je, že jednotlivé metódy monitora sú spolu úzko previazané, ich sémantika je zložitá, a celá zručnosť viacvláknového programovania v jazyku Java je založená na nastavení a kontrole tohto vzájomného pôsobenia [7]. Na rozdiel od toho, je paralelná kompozícia CSP procesov matematicky jasná, jednoznačná a škálovateľná s rastom systému.

### IV. ABP (ALTERNATING BIT PROTOCOL)

Podstatou riešenia je implementácia vybraného systému s paralelnými prvkami dvoma spôsobmi. S využitím knižnice JCSP a bez použitia tejto knižnice. Výsledné implementácie budú porovnané vzhľadom na zložitosť a efektívnosť daných implementácií. Predpokladom je, že knižnica JCSP by mala uľahčiť správu paralelných procesov.

Na porovnanie návrhu a implementácie paralelných programov s využitím knižnice JCSP a bez jej využitia bude použitá simulácia činnosti ABP (Alternating Bit Protocol). ABP je jednoduchý sieťový protokol pracujúci na spojovej vrstve OSI referenčného modelu. Je to protokol, ktorý zabezpečuje spoľahlivú komunikáciu cez médium, na ktorom môže dôjsť k strate správy, nie však k jej poškodeniu. Protokol pozostáva zo 4 procesov S, R, E a F. Proces S je odosielateľ a proces R je prijímateľ. Procesy S a R sú prepojené kanálmi E a F, prostredníctvom ktorých dochádza k výmene správ medzi odosielateľom a príjemcom [8]. Kanály E a F predstavujú komunikačné médium medzi procesmi, na ktorom môže dôjsť k strate odoslanej správy.



K dátam, ktoré odosielateľ  $S$  posiela príjemcovi  $R$ , sa pripoji informácia v podobe jedného bitu, buď  $0$  alebo  $1$ , ktorá slúži na potvrdenie úspešného prijatia správy. Kanálom  $E$  sú prenášané správy, ktoré pozostávajú z dát, ktoré  $S$  posiela  $R$ , a z jednobitovej informácie, tzv. potvrdenia. Kanálom  $F$  sú prenášané len potvrdenia, jednobitové správy, ktoré slúžia pre procesy na určenie úspešného, respektíve neúspešného odoslania a prijatia správy.

Procesy odosielateľa a príjemcu majú uložené kontrolné bity. Kontrolné bity sú porovnávané s potvrdzujúcimi bitmi prijatými v správe. Ak sa zhodujú, tak komunikácia prebehla úspešne. Ak sa nezhodujú, tak proces odosielateľa nezačne odosielať novú správu, ale pokračuje v posielaní danej poslednej správy [8] a proces prijímateľa nezapíše správu na výstupný kanál.

#### A. Návrh a implementácia s využitím JCSP

Na vytvorenie protokolu ABP na základne formálnych metód je potrebný popis a návrh systému pomocou niektorej z formálnych metód, v tomto prípade pomocou procesnej algebry CSP. Je to kľúčovou, najpodstatnejšou časťou pri vytváraní systému týmto spôsobom. Na implementáciu je použitý ABP model, ktorý opísal A. W. Roscoe vo svojom diele *The Theory and Practice of Concurrency* [8]. Nachádza sa tam takisto jeho formálny popis procesnou algebrou CSP, použitý v tejto práci, ktorý je zapísaný syntaxou pre nástroj FDR 2. Jedná sa o simuláciu činnosti ABP, a teda komunikačné médium  $E$  a  $F$  je navrhnuté tak, aby pri komunikácii prostredníctvom neho došlo k strate niektorých správ.

Postup vytvárania špecifikácie systému priamo v jazyku pre nástroj FDR je nasledovný:

- Špecifikujú sa údajové štruktúry a dátové typy, ktoré budú potrebné pri definovaní činnosti procesov.
- Určia sa kanály, prostredníctvom ktorých budú procesy komunikovať, a zároveň typ údajov, ktoré budú kanálmi posielané.
- Vytvorí sa špecifikácia systému, zjednodušene, ako sa bude systém javiť navonok, ako celok. V tomto prípade je špecifikácia podobná buffer-u s veľkosťou danou číslom  $N$ .

Po určení špecifikácie systému sa vytvorí jeho implementácia. Popíše, implementuje sa správanie jednotlivých procesov pomocou prvkov ponúkaných CSP algebrou. Z daných procesov sa následne zostaví implementácia systému. Protokol je určený na komunikáciu prostredníctvom kanálov, na ktorých dochádza k strate správ. Takéto kanály sú simulované procesmi  $E$  a  $F$ , ktoré zabezpečujú to, že môže dôjsť k strate prijatej správy, no nedôjde k strate nekonečnej postupnosti správ. Implementácia protokolu pozostáva z odosielajúceho procesu *Send* a prijímajúceho procesu *Receive*, ktoré sú prepojené pomocou  $E$  a  $F$ .

Takto zostavený systém je pripravený na proces verifikácie. Verifikácia je vykonaná pomocou nástroja FDR2 [9].

Formálne špecifikovaný a verifikovaný systém je pripravený na implementáciu v programovacom jazyku Java s využitím dostupnej knižnice JCSP, ktorá umožňuje implementáciu jednotlivých prvkov systému.

#### Implementácia navrhnutého systému

Implementácia procesov  $E$  a  $F$ , reprezentujúcich nespoľahlivé kanály, je v podstate rovnaká. Líši sa jedine v type dát, ktoré z kanálov čítajú, a ktoré na ne zapisujú. Ak počet neodoslaných (stratených) správ dosiahol limit  $L$ , správa je poslaná ďalej, ináč sa vykoná jedna z alternatív. Buď sa správa odošle ďalej a počítadlo je vynulované, alebo sa správa „zahodí“ a zaznamená sa neodoslania správy.

Proces odosielajúci správy, proces *Send*, sa vykonáva v nekonečnom cykle. Ak je hodnota správy *Null*, zo vstupu je načítaný nový reťazec a kontrolný alternujúci bit je zmenený. Ak hodnota správy nie je *Null*, teda zo vstupu bola načítaná správa, ktorá nebola odoslaná, tak sa vykoná jedná z alternatív. Buď je správa odoslaná ďalej, alebo je prijatý potvrdzujúci bit, ktorý ak sa zhoduje s kontrolným bitom procesu *Send*, je hodnota správy nastavená na *Null*, a tak proces môže zo vstupu načítať ďalší reťazec.

Proces prijímajúci správy, proces *Receive* sa vykonáva v nekonečnom cykle. Proces prečíta správu, ktorú mu odoslal proces *Send*. Ak je potvrdzujúci bit, pripojený k tejto správe, zhodný s kontrolným bitom procesu *Receive*, správa je zapísaná na výstup, a jej potvrdzujúci bit je odoslaný ako potvrdenie procesu *Send*. Kontrolný bit procesu je zmenený. Ak sa bity nerovnajú, dôjde len k odoslaniu potvrdzujúceho bitu prijatej správy procesu *Send*.

Zjednodušený zápis, určený len na ilustráciu vzťahu medzi procesmi, by mohol vyzeráť takto:

Send || ((E || F) || Receive)

Symbol „||“ znamená paralelnú kompozíciu procesov. Tento vzťah medzi procesmi je zachovaný aj pri jeho prepise do programovacieho jazyka.

### B. Návrh a implementácia bez využitia JCSP

Vytvorený program, bez využitia JCSP, je simuláciou činnosti ABP. Tvoria ho 4 samostatné vlákna, ktoré sú definované v triedach `Send2`, `Receive2`, `E2`, `F2`.

`E2`, `F2` ako vlákna reprezentujúce nespoľahlivé kanály, sú implementované tak, že v nekonečnom cykle čítajú zo svojho vstupu dáta, a na základe náhodne generovaného čísla sa určí, či dané dáta budú odoslané alebo budú zahodené. Týmto náhodným určením je simulovaná strata správ. Pri každom zahodení dát sa premenná  $n$  zníži o jedno. Ak  $n = 0$ , teda ak dáta boli  $n$ -krát po sebe zahodené, dáta sa odošlú.

`Send2` vo funkcii odosielateľa vykonáva v nekonečnom cykle niekoľko hlavných krokov. Ak nemá dostupnú správu, načíta ju z užívateľského vstupu a zmení kontrolný bit. Kontrolný bit pridá na koniec správy, kde plní funkciu potvrdzujúceho bitu. Správu následne odošle. Ak na kanál od procesu `F` bol zapísaný potvrdzujúci bit, prečíta ho. Ak sa zhoduje s kontrolným bitom, tak aktuálnu správu zmaže, čím v ďalšom cykle dôjde k načítaniu novej správy.

Prvotná implementácia triedy `Send2` bola testovaná s použitím kanálov, na ktorých nedochádzalo k strate správ. Činnosť vlákna `Send2` a aj celého systému bola správna a nedošlo k žiadnym problémom. No po simulácii straty správ pri ich posielaní kanálmi sa systém pri vykonávaní svojej činnosti takmer okamžite dostal do stavu vzájomného čakania vlákien, deadlock. Činnosť vlákna `Send2` bola pozastavená na akcii čítania z kanálu od procesu `F`, no kvôli strate správ nedošlo k zápisu žiadnej informácie na tento kanál. Preto akcia čítania z tohto kanálu nebola možná. Problém si vyžiadal dodatočnú synchronizáciu a usmernenie čítania z kanálov, a teda prídanie testovania, či došlo k zápisu na kanál.

`Receive2` plní funkciu príjemcu. V nekonečnom cykle sa vykonávajú nasledujúce hlavné kroky. Zavolá sa metóda, ktorá zabezpečí načítanie dát so vstupného kanálu. Ak sa potvrdzujúci bit, nachádzajúci sa na konci prijatej správy, zhoduje s kontrolným bitom tohto vlákna, správa sa vypíše na používateľský výstup a kontrolný bit je zmenený. Následne je odoslaný potvrdzujúci bit procesu `Send2`.

Vytvorená implementácia bola síce funkčná, ale doba od zadania správy až po jej vypísanie bola príliš dlhá v porovnaní s prvým variantom systému. Vlákno zablokované na akcii čítania z kanálu, ktorý bol v tom čase prázdny, sa nevrátilo k vykonaniu akcie čítania okamžite potom, čo na kanál boli poslané dáta. Práve preto bola za každou akciou zápisu dát na kanál zavolaná metóda `flush()` nad daným objektom reprezentujúcim kanál. Metóda `flush()` zabezpečí, že každý objekt, ktorý čaká na čítanie z daného kanálu, bude upozornený, že môže čítať a zabezpečí, že dáta z kanálu budú hneď prečítané. V niektorých prípadoch došlo až k viac ako 10 násobnému zrýchleniu vykonania úlohy. Priemerný čas vykonania úlohy sa s použitím metódy `flush()` 6 násobne zmenšil z hodnoty 2,14 ms na hodnotu 0,37 ms.

## V. POROVNANIE RIEŠENÍ

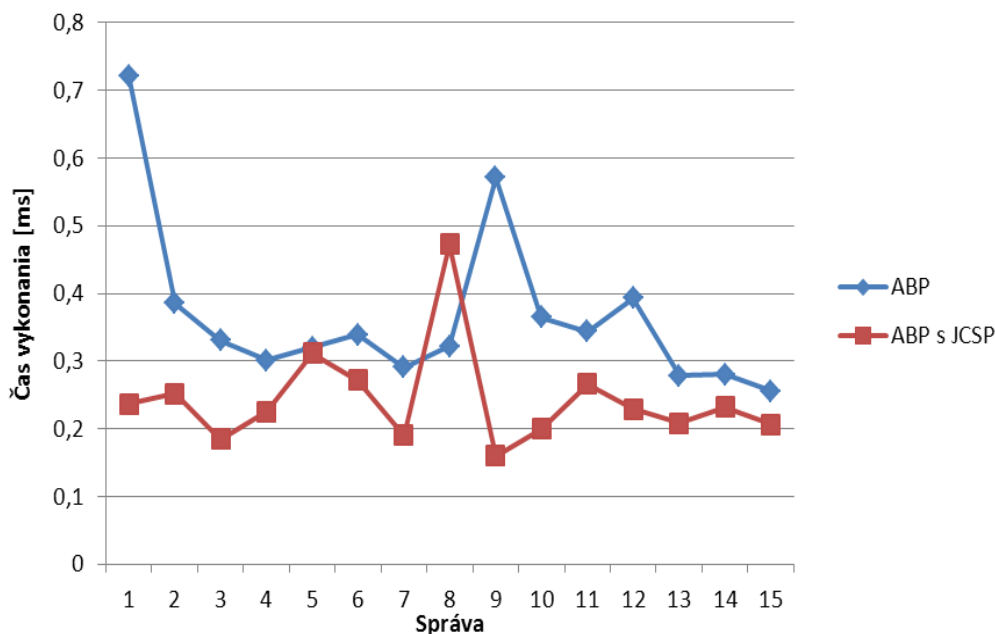
Riešenia, dve implementácie toho istého systému, simulácie ABP, sú porovnané z viacerých hľadísk.

### A. Rýchlosť vykonania úlohy

Prvým z nich je rýchlosť vykonania úlohy, na ktorú je protokol určený. Bolo vykonaných 15 meraní času od načítania reťazca z používateľského vstupu po jeho vypísanie na štandardný výstup. To znamená, že bolo odoslaných, a samozrejme aj prijatých, 15 správ prostredníctvom jednej aj druhej implementácie ABP. Výsledky 15 meraní dĺžky trvania vykonania operácie odoslania reťazca, pre obidve implementácie sú uvedené v grafe Obr. 1.

Z výsledkov merania vyplýva, že úloha, ktorú má plniť ABP, je vykonaná v kratšom čase v implementácii ABP s využitím JCSP. Čas vykonania je v priemere o 33,4% kratší, ako čas vykonania úlohy v implementácii bez použitia JCSP.

Rýchlejšie vykonanie danej úlohy implementáciou s použitím JCSP bolo zabezpečené hlavne vďaka princípom CSP. Primitíva, ktoré ponúka na komunikáciu medzi procesmi, zabezpečujú rýchlejšiu komunikáciu medzi procesmi, keďže tie sú synchronizované práve na zápise a čítaní z kanálov. Tieto dve udalosti nastávajú súčasne.



Obr. 1 Graf trvania odoslania správy v jednotlivých implementáciách

### B. Počet riadkov zdrojového kódu

Ďalším hľadiskom porovnania dvoch implementácií, je ich dĺžka, počet riadkov zdrojového kódu jednotlivých implementácií, jednotlivých tried.

Riadky obsahujúce definície premenných, len uzatvárajúce zátvorky, prázdne riadky a riadky obsahujúce len komentáre sú vynechané.

Väčšina tried implementácie s využitím JCSP, a takisto implementácia ako celok, má viac riadkov zdrojového kódu ako druhý variant ABP. Celkový počet riadkov implementácie s využitím knižnice JCSP je len o približne 5% väčší ako počet riadkov implementácie bez jej využitia. Vzhľadom na to, že rozdiel medzi dĺžkou implementácií je pomerne malý, je možné tvrdiť, že implementácie sú si vzhľadom na dĺžku veľmi podobné, skoro až rovnaké.

### C. Zložitosť návrhu a implementácie

Tretím hľadiskom porovnania je zložitosť návrhu a následnej implementácie jednotlivých variantov.

Proces formálnej špecifikácie systému nie je jednoduchý. Je potrebné poznať procesnú algebru CSP a jej jednotlivé stavebné prvky a dokázať pomocou nej definovať systém. Systém definovaný pomocou CSP algebry je nutné prepísať pomocou jazyka, ktorý používa overovací nástroj FDR2 [9]. Následne pomocou nástroja FDR2 vykonať verifikáciu systému, ktorá overí neprítomnosť stavov deadlock a livelock v systéme. Práve definícia systému a jeho verifikácia tvoria najzložitejšiu ale aj najdôležitejšiu časť návrhu systému. Samotná implementácia systému je jednoznačná a pomerne jednoduchá. Po oboznámení sa so základnými triedami balíka JCSP je možné ihneď začať s implementáciou. Po správnom interpretovaní formálnej definície jednotlivých procesov je systém funkčný a nie je potrebné ho ďalej testovať, či sa nedostane do stavu deadlock alebo livelock.

Druhý variant vytvorenia systému si nevyžaduje formálny popis systému. Návrh tohto variantu systému predstavoval len popis činnosti protokolu ABP. Kým návrh si nevyžaduje toľko úsilia ako pri formálnom popise a je jednoduchý, tak samotná implementácia je podstatne náročnejšia ako pri prvom variante vytvárania systému. Programovací jazyk Java ponúka viacero možností pri implementácii jednotlivých prvkov, čo nemusí byť vždy len pozitívne, v tomto prípade to bolo skôr mátuče. Pri implementácii systému bolo nutné riešiť problémy, ako napríklad synchronizácia vlákien. Spracovanie správy ďalším vláknom nenastalo okamžite po jej odoslaní, preto bolo potrebné pridať oznámenie vláknam, že údaje na ktoré čakajú, boli zapísané a môžu ich čítať. Aj keď test prebehne úspešne, nie je zaručené že hazard nemôže nastať, teda že sa systém nedostane do nežiaduceho stavu. Taktiež nie je zaručené, že vytvorený systém zodpovedá požiadavkám, a že pracuje správne.

## VI. ZÁVER

Výsledkom porovnania dvoch spôsobov implementácie toho istého systému je, že pri jednoduchých paralelných systémoch, akým je aj protokol ABP, je jednoduchšie nepoužiť formálny prístup k riešeniu, a s tým späť knižnicu JCSP. Porovnateľný výsledok je, samozrejme, možné dosiahnuť aj prostredníctvom programovacieho jazyka Java, bez špeciálnych knižníc, akou je JCSP. Ak je ale dôraz kladený na spoľahlivosť a takisto na rýchlosť komunikácie procesov systému, je formálny prístup vhodnejším riešením.

Závery vytvorené na základe porovnaní dvoch variantov implementácie protokolu ABP platia pre daný prípad, jednoduchý paralelný systém. Nemusia nutne platiť v rovnakej miere pre každý paralelný systém vytvorený týmito dvoma spôsobmi. Avšak vzhľadom na komplikácie pri implementácii jednoduchého komunikačného protokolu, je možné predpokladať, že pre systémy s väčšou zložitou a rozsahom bude nutné pre zaručenie ich spoľahlivosti použiť formálny prístup. Taktiež možno predpokladať, že zložitejšie a väčšie paralelné systémy bude jednoduchšie vytvoriť a následne aj spravovať, ak sa pri ich vytváraní použijú formálne metódy, napríklad CSP. Dosiagnuté výsledky sú vhodným podnetom pre ďalšie pokračovanie výskumu v oblasti riešenia problému vývoja spoľahlivých a bezpečných paralelných systémov.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (ITMS 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] COLLINS, Michael: Formal Methods [online]. Dependable Embedded Systems, 1998. [cit 2013-10-29] Dostupné na internete: <[http://www.ece.cmu.edu/~koopman/des\\_s99/formal\\_methods/](http://www.ece.cmu.edu/~koopman/des_s99/formal_methods/)>.
- [2] BAETEN J.C.M.: A Brief History of Process Algebra. In: Theoretical Computer Science - Process algebra Volume 335 Issue 2-3. 2005. il: ISSN 0304-3975.
- [3] WELCH, Peter H.: Parallel and Distributed Computing in Education (Invited Talk). Portugal: Third International Conference, 1998. ISBN: 978-3-540-66228-0.
- [4] HOARE, C.A.R.: Communicating Sequential Processes [online]. 2004. [cit. 2013-12-20]. Dostupné na internete: <<http://www.usingcsp.com/cspbook.pdf>>.
- [5] WELCH, Peter H.: Process Oriented Design for Java: Concurrency for All. London: Springer-Verlag, 2002. ISBN: 3-540-43593-X.
- [6] WELCH, Peter H. - AUSTIN, P.D.: CSP for Java API Specification [online]. [s.a.]. [cit. 2013-10-29]. Dostupné na internete: <<http://www.cs.kent.ac.uk/projects/ofa/jcsp/jcsp-1.1-rc4/jcsp-doc/>>.
- [7] WELCH, Peter - BROWN, Neil: Communicating Sequential Processes for Java [online]. 2011. [Aktualizované 2011-7-14]. [cit. 2013-10-29]. Dostupné na internete: <<http://www.cs.kent.ac.uk/projects/ofa/jcsp/explain.html>>.
- [8] ROSCOE, A. W.: The Theory and Practice of Concurrency. Upper Saddle River, New Jersey: Prentice Hall PTR, 1997. 512 s. ISBN 0136744095.
- [9] ROSCOE, Bill, ARMSTRONG, Philip et al.: FDR2 User Manual. [online]. 2009-2010. [s.a.]. [cit. 2014-03-28]. Dostupné na internete: <<http://www.fsel.com/documentation/fdr2/html/fdr2manual.html#Top>>.

# Processing of Measured Data with a Focus on Visibility

<sup>1</sup>Matúš TATARKO, <sup>1</sup>Luboš OVSENÍK, <sup>1</sup>Ján TURÁN

<sup>1</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>matus.tatarko@tuke.sk, <sup>1</sup>lubos.ovsenik@tuke.sk, <sup>1</sup>jan.turan@tuke.sk

**Abstract** — Proposal paper contains results of FSO availability estimation based on measurements on TUKE experimental system. The FSO systems use infrared beams to transmit data via atmosphere. Transmission channel is unstable and unpredictable, so monitoring of environment is needed. Atmosphere phenomena decrease optical signal in different ways such as absorption, scattering and scintillation. The main negative influence on laser beam is caused by small particles of water; dust and smog in the air which size corresponds with using wavelengths in FSO systems. Thus, it is important to know behavior of atmosphere in area where the FSO system is working. In this article are shown two methods how to measure and evaluate information about weather conditions in local area and estimate FSO availability from this data. Experimental results from long time measurements are also presented.

**Keywords** — availability, environment, free space optic, steady model

## I. INTRODUCTION

The FSO systems achieve high bandwidth communication between two static points via air. Transmitted data are modulated on infrared optical beam. The optical carrier frequencies are in the range of 300 THz [1]. These frequencies correspond with wavelengths in range of 850-1550 nm. Three atmospheric transmission windows are in this range, where the attenuation of atmosphere is minimal [2,3,4]. Significant negative influences on transmitted optical signal are caused by absorption, scattering and scintillation. These three effects occur in environment with the fog and low clouds. Atmospheric attenuators degrade the system performance. However other factors like rain and snow are generally less significant. Turbulence induced atmospheric scintillation causes severe fluctuation in received signal power [5]. Scintillation effects for terrestrial FSO links can be mitigated by increasing transmission power, wavelength diversity, multiple transmit beam and multiple receiver [6,7]. The main negative factor for terrestrial FSO links is fog and with fog closely related visibility.

In the second chapter is described fog effect on FSO system. The third part of this article is devoted to description of measuring visibility and its evaluation. Section four describes FSO system FSONA, SonaBeam 155E. In the last part of paper is calculated overall availability of the experimental FSO system at the Technical University of Košice.

## II. FOG EFFECTS ON THE FSO SYSTEMS

Fog can be characterized by several physical parameters such as liquid water content (LWC), particle size distribution, temperature and humidity. Fog composes of fine water droplets, ice crystals or smoke particles suspended in the atmosphere. Fog near the surface reduces the visibility and causes degradation of optical link performance. In general, the probability of occurrence of fog is much higher in winter than in summer for continental environments when temperature approaches 0 °C and relative humidity rises above 80 % [8].

Fog significantly reduces the availability and reliability of FSO links. The attenuation due to fog in maritime environment reaches up to 480 dB/km and 236 dB/km for continental fog condition [8]. According to international code of visibility, there are four types of fog depending on visibility range in Table 1.

Table 1  
 Types of Fog and Visibility of Relevant

Visibility range (m)	Type of fog	Attenuation (dB/km)
40 - 70	Dense fog	250 - 143
70 - 250	Thick fog	143 - 40
250 - 500	Moderate fog	40 - 20
500 - 1000	Light fog	20 - 9,3

### III. MEASURING OF VISIBILITY

Visibility was first defined for meteorological purposes as a quantity to be estimated by a human observer, and observations made in that way are widely used. Overall, there are three instruments for measuring the extinction coefficient; Telephotometric instruments, Visual extinction meters and Transmissometers.

The transmissometer consists of a laser transmitter and a receiver. They are separated by a certain distance (maybe 10 m). The transmitter sends a laser beam through open air toward the receiver. The instrument calculates how much the beam is attenuated by the atmospheric phenomenon and calculates visibility. Some weather station installations will also have an instrument called a forward scatter visibility sensor. It works in a similar fashion to the transmissometer, except it uses infrared light rather than a visible laser. The transmitter and the receiver are close to each other and are offset at an angle. If there is a visibility-reducing phenomenon present, the infrared light will be scattered by particles suspended in the air. The offset angle allows the receiver to measure how much of the light has been scattered, and from that measurement it can calculate the visibility. Transmissometers are more accurate in lower visibility conditions, and forward scatter visibility sensors are better for accurately measuring clearer conditions.

#### A. Information about visibility in Europe

Measured information about visibility in Europe is able on weather forecast websites which give us every day values of visibility (km). One of these websites is <http://www.wunderground.com/global/Region/EU/Visibility.html>. If we need long time information about visibility, there is only a color map which gives us only informative value of visibility Fig. 1.

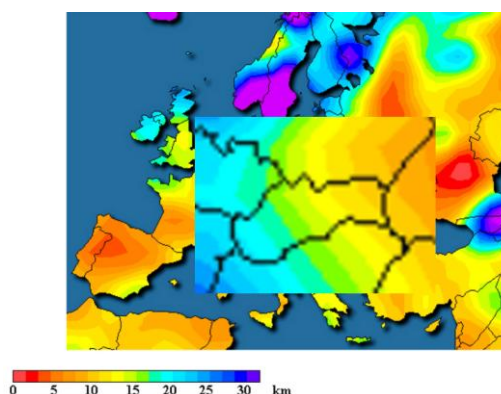


Fig. 1. Color map of visibility for Slovakia

In this figure we can see only color scale with some values of visibilities which belongs to a color map. But if we need exact value of visibility in certain city, there is a problem. Almost every weather forecast websites does not provide information about visibility, because devices for measuring visibility are expensive. Thus only big cities or airports have this information.

As we mentioned above, devices for direct measuring of visibility are expensive, so there is a way how to calculate visibility from another parameter of fog. It is called Liquid Water Content (LWC). It is one of the most important parameters of the fog. LWC describes the mass of water drops in the volume units of the fog. It is measured in  $[g/m^3]$ . Direct method for measuring LWC consists of extracting a known volume trough a cotton pad or of rotating cups in an impeller apparatus, both to the weighed. Also resistance changes can be measured with a hot wire probe attached to an aircraft flying through the clouds. The value of LWC in the fog varies

in a wide range. In a most of the fog attenuation models, it is required to know the value of LWC [9].

### B. Fog sensor

Fog sensor device was developed as a low-cost device for measuring atmospheric parameters such as temperature [°C], relative humidity [%] and density of fog [g/m<sup>3</sup>] in local area [10,11]. They are important parameters for subsequent static and statistical evaluation of the environment quality of free space communication. Fog sensor has two main parts (Fig.2). External unit performs a measurement and internal unit convert physical interface from RS 422 line to RS 232 line. Indoor unit is connected to PC and data are stored on server in .txt files.

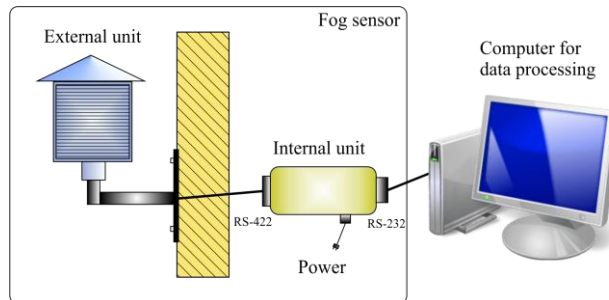


Fig. 2. Main parts of Fog sensor

Sensor for measuring density of fog is stored in a box which protects it from direct sunlight and rain, but allows free movement of particles to the sensor. Sensor sends and receives infrared pulses which are attenuated on water and dust particles (Fig.3).

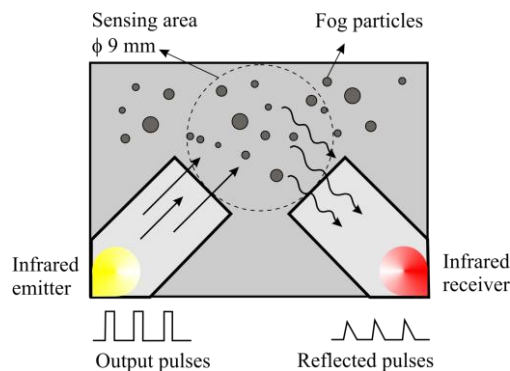


Fig. 3. Principle of measuring density of fog.

### IV. FSONA, SONABEAM 155E

The FSO system FSONA, Sonabeam 155E is an extremely reliable, very high speed, full-duplex, wireless component for communication network employing multiple, infrared, eye safe laser transmitters, a large aperture optical receiver and an embedded microprocessor. Terminals are designed to operate equally well either indoors or outdoors where a clear and unobstructed line of sight exists between two static connecting points in a network. Link distance can varies from 50 m to several km. Terminal Sonabeam 155E can be used to transport 1,5 - 160 Mbps with using eye safe 1550 nm lasers [12].

Table 2  
 Parameters of Sonabeam 155E

Parameters	FSONA, Sonabeam 155E
Wavelength	1550 nm
Tx Power	100 mW
Rx Power	-30 dBm
Rx Diameter	10 cm
Directivity	2 mrad



In our case we mention the connection between two buildings in campus of Technical University of Košice (TUKE) with distance of 451 m.

## V. AVAILABILITY OF FSO AND RESULTS

### A. Results of availability obtained by Fog sensor

The availability of FSO link we can obtain from measured data from fog sensor. The first step is to obtain threshold (m) in which the FSO link still works. For this, we use software application, which we developed at the TUKE, called FSO System Simulator, Fig. 4. There we can select system from database or we can enter input different device properties.

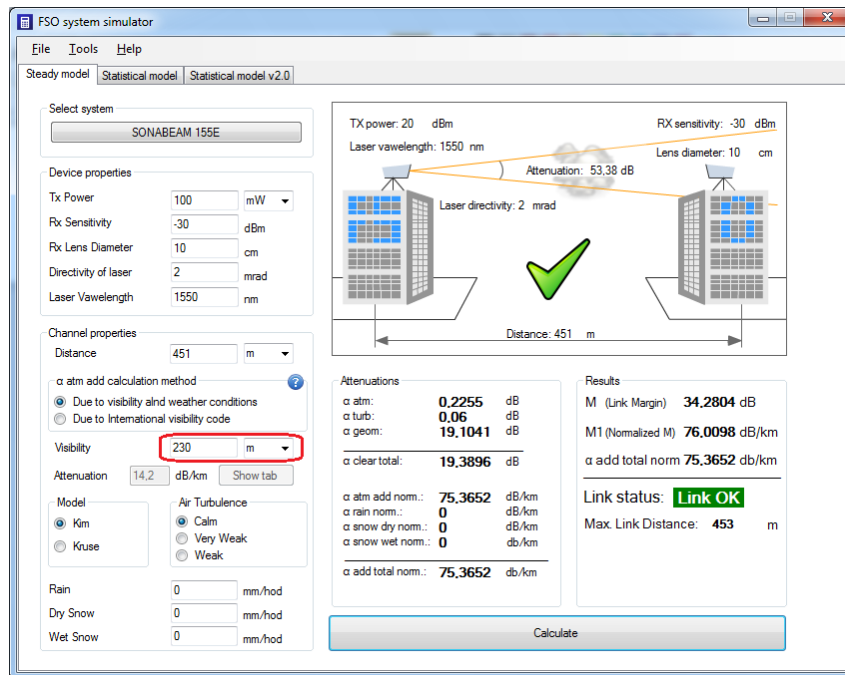


Fig. 4. The threshold for next calculation.

From Fig. 4 we can see that the threshold of visibility for 451m link is 230m. So from range of 0- 230m the FSO system, FSONA, Sonabeam 155E will be down. For next calculation we need to convert 230m to value of LWC, because fog sensor measures density of fog in g/m<sup>3</sup>. From LWC we can calculate the visibility and reverse. For the calculation we use the empirical formula for fog visibility as a function of fog density:

$$LWC = 2 \cdot C \cdot \rho \cdot V^{-1} \left[ g / m^3 \right] \quad (1)$$

where  $V$  is the visibility [m],  $\rho$  the droplet radius in microns and  $C$  a constant. The average water content in fog of various types may be computed from measured values of visibility and drop size with  $C= 2,6$ . Cloud, fog and mist droplets are very small. Their mean diameter is typically 10-15 micron [13,14]. In our case, the relation for LWC is:

$$LWC = \frac{2 \cdot 2,6 \cdot 15}{231} = 0,2947 \left[ g / m^3 \right] \quad (2)$$

This obtains value we use in the last calculation in FSO System Simulator. Third card of this application allows us enter input device parameters, time span of occurrence fog event and the threshold value. In this case, it is 0,2947 g/m<sup>3</sup>, Fig. 5.

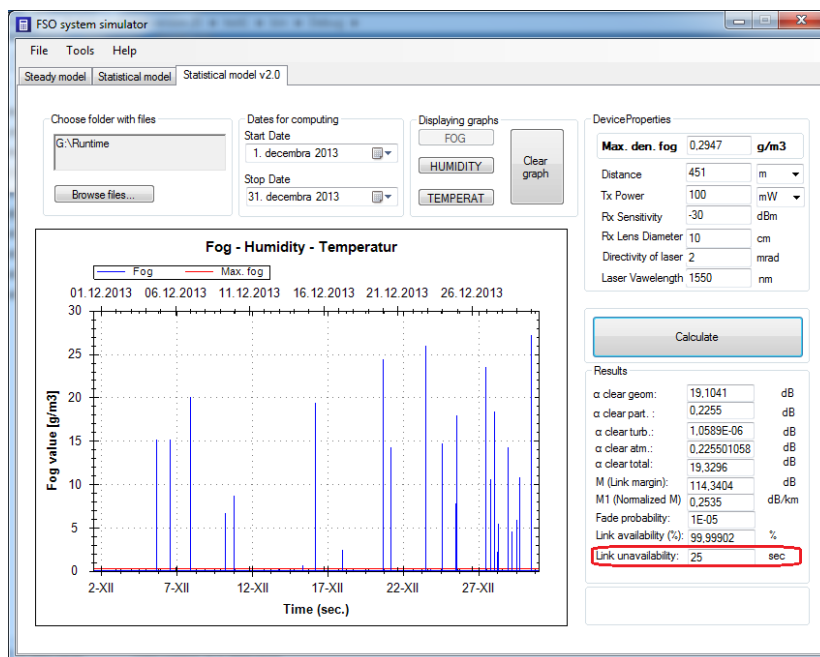


Fig. 5. Calculated unavailability of FSO system during whole December 2013.

All values, which are higher than threshold cause unavailability of FSO link and they are registered in separate .txt files, Fig. 6.

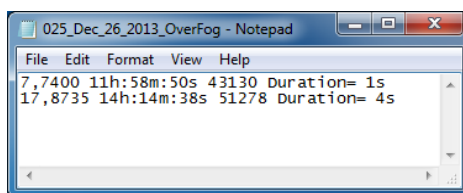


Fig. 6. Text file with recorded data.

In this case, is able to see, that link unavailability during December 2013 was 25 seconds. That represents 99, 99902 % availability of FSO link during this month.

### B. Results of availability obtained by SHMI

Second type of calculation is based on information from Slovak Hydrometeorological Institute (SHMI). Data was obtained in Košice at hourly intervals. Each outage is one hour, although the outage could be shorter. During 4 years i.e. 2009-2012 were detected 287 outages (hours). So, availability of FSONA link was 99, 18094 %.

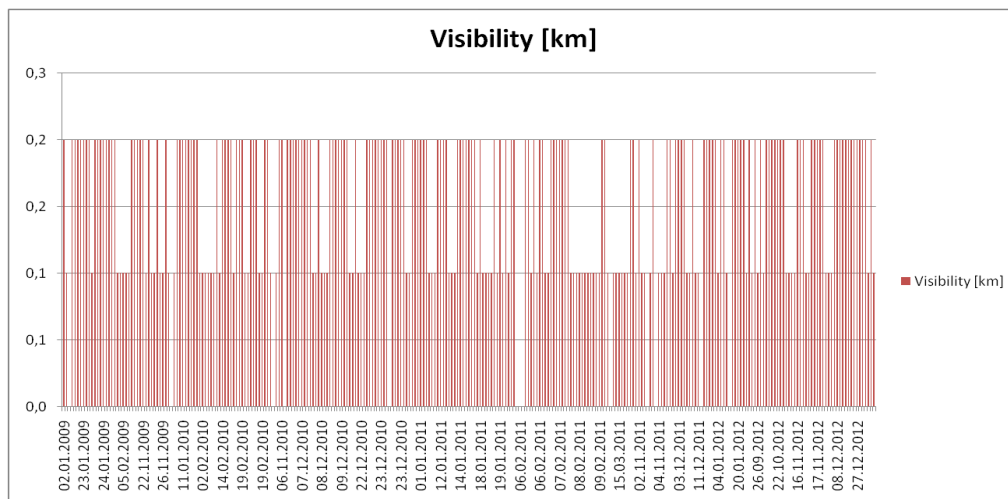


Fig. 7. Lower visibility than link distance during four years 2009- 2012.

## VI. CONCLUSION

In this article were described two methods for estimation of experimental FSO system availability. The FSO link with FSONA transceivers was tested on 451 m at the campus of TUKE. First method obtains data from Fog sensor in seconds' intervals. Each of outages was defined as one or more seconds (e.g. 4 sec.). Second method obtains data from SHMI in hours intervals. Each of outages was defined as one hour outage. In results is graphical shown availability of FSO link. From mentioned above we can see that first method with Fog sensor gives better results. The main disadvantage of second method is period of obtaining data (1 hour in comparison to 1 second in first method).

## ACKNOWLEDGMENT

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „063TUKE-4/2013 - The Use of Remote Controlled Optical Fibre Refractometer in Teaching“, and project no. „006TUKE-4/2014 - The Use of TUKE PON Experimental Model in Teaching. This work is also the result of the contract no. “APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks” supported by the Slovak Research and Development Agency.

## REFERENCES

- [1] L. Ovseník, J. Turán, P. Mišenčík, J. Bitó, L. Csurgai-Horváth, “Fog Density Measuring System,” *Acta Electrotechnica et Informatica*, 2012, pp. 1-5.
- [2] R. Kvicala, V. Kvicera, M. Grabner, O. Fiser, “BER and Availability Measured on FSO Link,” *Radioengineering*, Vol. 16, No. 3, 2007.
- [3] P. Hovořák, O. Wilfert, “Free space optics- Measurement of Transmission Quality Link Parameters,” *Electronics'2005*, 21- 25 September, Sozopol, Bulgaria, 2005.
- [4] M. O. Zafari, “Wireless Optical Communication Systems in Enterprise Network,” *Telecommunication Review*, 2003.
- [5] B. R. Strickland, M. J. Lavan, E. Woodbridge, V. Chan, “Effect of Fog on the Bit-error Rate of a Free-space Laser Communication System,” *Applied Optics*, Vol. 38, pp. 12-19, 1999.
- [6] D. Giggenbach, H. Henniger, N. Perlot, F. David, “Multiplewavelength Free-space Laser Communications,” *In Proc. of SPIE, San Jose, CA, USA*, Vol. 4975, pp. 12-19, 2003.
- [7] A. Biswas, S. Lee, “Ground-to-Ground Optical Communications Demonstration,” *Jet Propulsion Laboratory, Pasadena, CA, Telecommunications and Mission Operation Progress, TMO Progress Report 42-141*, pp. 1-31, 2000.
- [8] S. S. Mohammad, B. Flecker, E. Leitgeb, M. Gebhart, “Characterization of Fog Attenuation in Terrestrial Free Space Optical Links,” *Journal of Optical Engineering*, Vol. 46, No. 6, pp. 1-10, 2007.
- [9] M. Reymann, J. Piasecki, F. Hosein and col. *Meteorological Techniques*, July, 1998.
- [10] L. Csurgai-Horváth, J. Bitó, “Fog Attenuation on V Band Terrestrial Radio and a Low Cost Measurement Setup,” *Future Network & Mobile Summit, Florence, Italy, Paper #47*, June 2010.
- [11] M. Tatarko, L. Ovseník, J. Turán, “Availability and Reliability of FSO Links Estimated from Visibility,” *Carpathian Journal of Electronic and Computer Engineering*, Vol. 5, pp. 121-126, 2012.
- [12] Sonabeam E-series Installation manual, Document No. 95-0224, July, Canada, 2006.
- [13] H. J. Kampe, “Visibility and liquid water content in clouds in the free atmosphere,” *Journal of meteorology*, September, 1949.
- [14] W. Steingartner, D. Galinec, “The Rôle of Categorical Structures in Infinitesimal Calculus,” *Journal of Applied Mathematics and Computational Mechanics*, Vol. 12, No. 1, pp. 107-119, 2013.

# Experimentálny model PON siete pre areál TUKE

<sup>1</sup>Matúš TATARKO, <sup>1</sup>Tomáš IVANIGA, <sup>1</sup>Luboš OVSENÍK, <sup>1</sup>Ján TURÁN

<sup>1</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická Univerzita v Košiciach, Slovenská republika

<sup>1</sup>matus.tatarko@tuke.sk, <sup>1</sup>tomas.ivaniga@student.tuke.sk, <sup>1</sup>lubos.ovsenik@tuke.sk,  
<sup>1</sup>jan.turan@tuke.sk

**Abstrakt** — Tento príspevok je zameraný na opis programového balíka OptSim a možnosti merania optických vlákien metódou OTDR. Simulačný softvér OptSim je profesionálny nástroj určený na simuláciu rôznych typov optických a opticko-elektrických komunikačných systémov. Jeho najväčšou výhodou je, že umožňuje simuláciu optických sietí pomocou rôznych režimov a módov simulácii, ktoré buď do simulácie zahrnú rôzne nelinearity a šumy alebo nie. OTDR poskytuje informáciu o celkovej kvalite optického vlákna, umožňuje zisťovať pozdĺžnu homogenitu optického vlákna, útlm zvarov a konektorov, dĺžku vlákna a zároveň lokalizovať jednotlivé prvky, prípadne poruchy optickej siete. V úvodnej kapitole je vysvetlený OptSim a je vytvorená optická sieť, ktorá predstavuje dve budovy spojené optickými vláknami. Druhá kapitola sa zameriava na OTDR, je vysvetlený Rayleighov rozptyl, Fresnelov odraz, mŕtva zóna a sú odmerané dve budovy v rámci areálu Technickej Univerzity v Košiciach (TUKE). V závere sú popísané jednotlivé výsledky získané z OptSimu a OTDR.

**Kľúčové slová** — Meranie, OTDR, optická sieť

## I. ÚVOD

Projektovanie optických vláknových komunikačných systémov sa zameriava hlavne na minimalizáciu prenosových strát. Najväčšou výhodou softvérových nástrojov je, že dovoľujú vo svojom vývojovom prostredí navrhnuť požadovaný komunikačný systém a plne ho odsimulovať na množstvo nežiaducich vplyvov a javov, tak aby čo najviac vyhovoval požiadavkám zadávateľa. Najväčšou chybou pri výstavbe komunikačných systémov je, že sa nevenuje dostatok času softvérovej analýze siete a pri fyzickej realizácii sa následne zistí, že daný systém pracuje inak ako sa predpokladalo. Výkopové práce sú finančne najnákladnejšou záležitosťou pri celom projekte a vykonávať rekonštrukciu novej siete nie je príjemná záležitosť [1]. V dnešnej dobe sa často stretávame s potrebou zistiť, v prípade poškodeného alebo prerušeného optického vlákna, v akej vzdialenosti od jeho začiatku sa táto porucha nachádza. Metóda spätného rozptylu nazývaná tiež optická časová reflektometria (Optical Time Domain Reflectometer – OTDR) je založená na periodickom vysielaní krátkych optických impulzov do optického vlákna. OTDR funguje prostredníctvom meracej metódy spätného odrazu, kedy je krátky optický impulz vyslaný optickým vláknom [2].

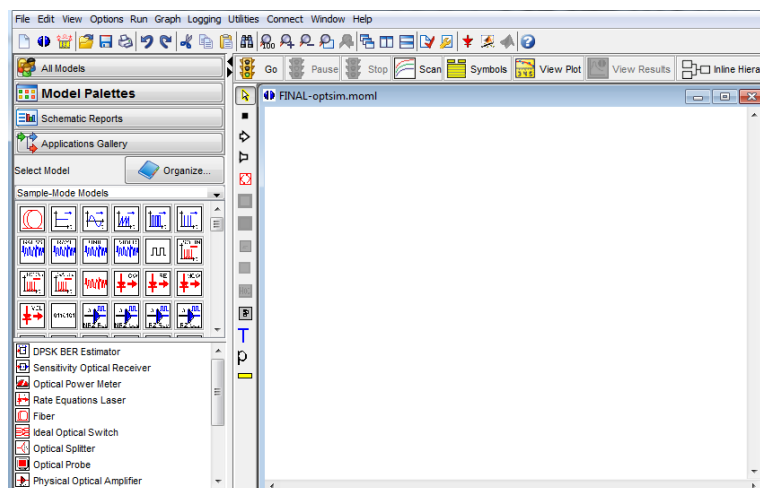
## II. PROGRAMOVÝ BALÍK OPTSIM

Simulačný program OptSim ponúka bohaté možnosti na simuláciu rôznych technológií. Vďaka svojmu prehľadnému grafickému rozhraniu, Obr.1 umožňuje vytvárať scenáre pre rôzne simulácie [3].

Programový balík OptSim je pokročilý simulačný program, určený pre profesionálny výskum v oblasti rôznych typov multiplexov (WDM, DWDM, TDM), káblovej televízie (CATV), optických sietí LAN, paralelných optických zberníc a ďalších optických systémov v oblasti telekomunikácii, ako aj dátových komunikácii pracujúcich s jednovidovými optickými vláknami [4].

Značnú výhodu prináša možnosť navrhnutia topológie v blokoch alebo po vzorkách, ktorú je možné prepojiť s kódom z Matlabu pre vlastný prvok alebo spracovanie. Aj pre neznalého

používateľa, ktorý sa v programe neorientuje, ponúka OptSim tzv. ukážkové modely, ako vo vzorkovom režime, tak aj v blokovom režime.

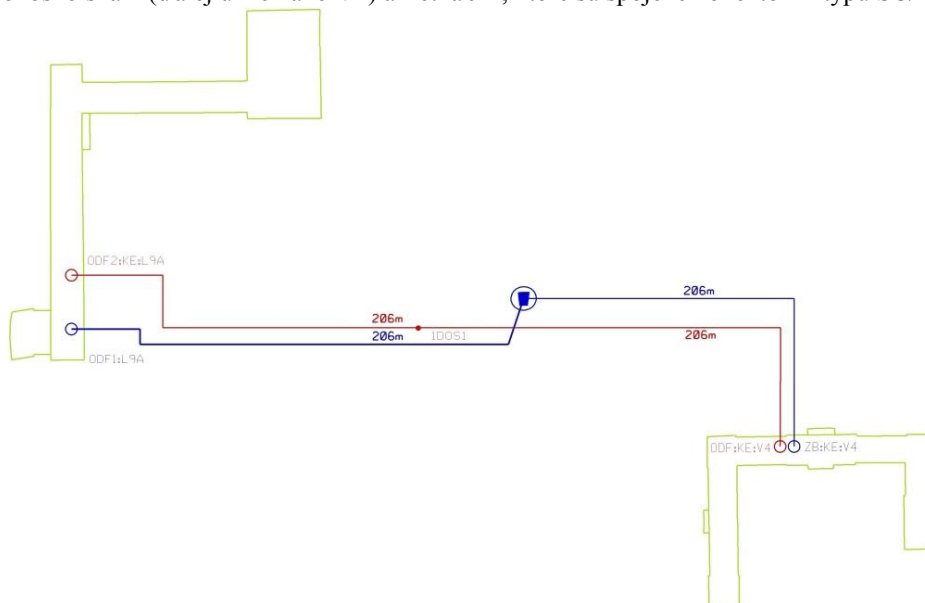


Obr. 1 Grafické rozhranie programu OptSim.

Návrh optickej siete v programe OptSim umožňuje exemplárne preveriť navrhnutú optickú sieť pred jej výstavbou. Taktiež umožňuje demonštratívne pripojenie optických meracích prístrojov a zobrazenie ich výsledkov: chybovosť BER, diagram oka, otvorenosť oka, medzisymbolovú interferenciu, výkonové úrovne, optické spektrum, elektrické spektrum a ďalšie. Nevýhodou aplikácie je absencia obojsmerného simulovania (obojsmerné simulovanie možno čiastočne nahradiť privedením spätných slučiek). Simulácia prebieha iba blok po bloku, kedy sú dáta z jedného bloku odovzdané ďalšiemu bloku [3,4].

#### A. Vytvorenie experimentálneho merania medzi dvoma budovami

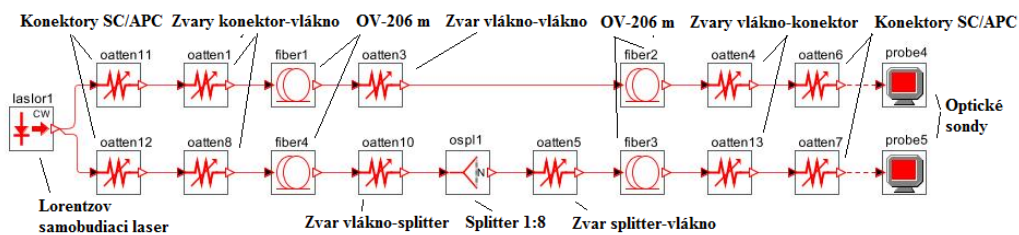
Pre vytvorenie experimentálneho merania je použitý zjednodušený vláknový plán, ktorý reprezentuje dve budovy Technickej univerzity v Košiciach. Na Obr.2 sú znázornené budovy Vysokoškolská 4 (ďalej už len ako V4) a Letná 9A, ktoré sú spojené konektormi typu SC/APC.



Obr. 2 Zjednodušený vláknový plán reprezentujúci dve budovy v areáli TUKE.

Pri meraní bola použitá optická sonda, ktorá zobrazuje diagramy ako výkon signálu, fázu, okamžitú frekvenciu, spektrum optického signálu a dovoľuje niekoľko druhov analýzy ako je detekcia kanálov a hodnotenie OSNR. V sieti sa nachádza aj vláknový rozbočovač, čo je vlastne pasívny sieťový prvok, ktorý je určený výhradne v sieťach PON. Jeho funkciou je rozdeliť optický signál zo vstupu na niekoľko výstupov a v spätnom smere ho zlučovať.

Na Obr.3 je schéma vytvorená vo vzorkovom móde s jednotlivým popisom daných komponentov.

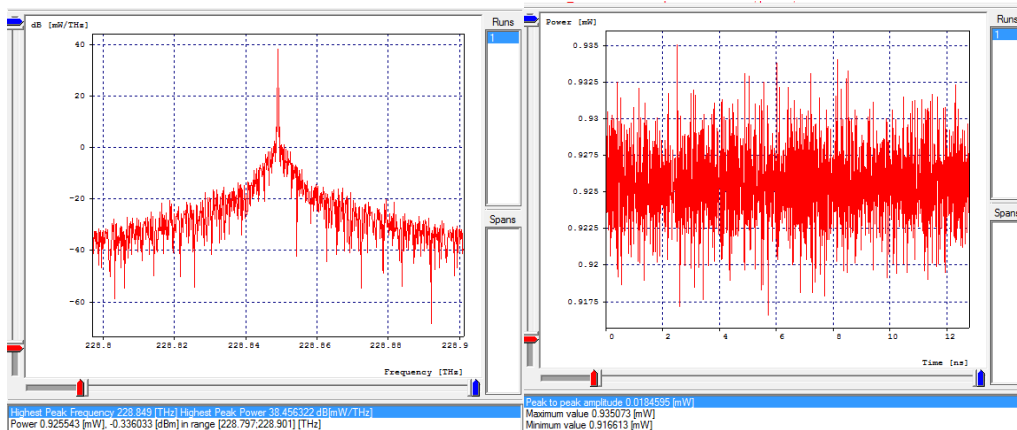


Obr. 3 Optická sieť vytvorená v programe OptSim.

### B. Simulácia s výsledkami

Na Obr.4 je optické spektrum a optický výkon na výstupe V4 pri vlnovej dĺžke 1310nm. Na konektore zapojenom v budove V4 je zaznamenaný výstupný výkon 0,925543 mW, ktorému prislúchal útlm 0,336033 dB. Daný optický výkon sa rozprestieral v oblasti od 228,797 až do 228,901 THz, pričom najvyšší výkon špičky (peaku) bol 38,456322 dB (mW/THz), ktorý bol zaznamenaný pri frekvencii 228,849 THz.

Maximálny výkon bol 0,935073 mW a minimálny 0,916613 mW. Hodnota výkonu špička - špička bola 0,0184595 mW.



Obr. 3 Optické spektrum a optický výkon na výstupe V4 pri vlnovej dĺžke 1330 nm.

Namerané výsledky v programovom balíku OptSim sú zaznamenané v Tab.1:

Tab. 1  
 Optické spektrum a optický výkon

Optické spektrum na výstupe V4		Optický výkon na výstupe V4	
Vlnová dĺžka [nm]	Útlm [dB]:	Vlnová dĺžka[nm]	min.-max. výkon [mW]
1310	0.336033	1310	0.916613-0.935073
1550	0.296064	1550	0.911992-0.957628
Optické spektrum so splitterom na výstupe V4		Optický výkon so splitterom na výstupe V4	
Vlnová dĺžka [nm]	Útlm [dB]	Vlnová dĺžka[nm]	min. -max. výkon[mW]
1310	11,1357	1310	0,0769225-0,0770583
1550	10,996	1550	0,0782274-0,080818

### III. OTDR

Pomocou tejto metódy na meranie prenosových parametrov optických trás je možné merať aj prenosové parametre jednotlivých prvkov trasy. Pri tejto metóde sa vyhodnocuje časová závislosť spätné rozptýleného optického výkonu pri šírení úzkeho optického impulzu meranou optickou trasou. Výsledkom je krivka spätného rozptylu v logaritmickej mierke, pričom vložený útlm konektora sa automaticky počíta z rozdielu výkonov signálu na vstupe a na výstupe tohto konektora. Útlm odrazu je daný integráciou plochy pod krivkou spätného rozptylu zodpovedajúcemu odrazu na príslušnom konektore. Táto metóda je dosť rozšírená v praxi [1].

Každý optický konektor je definovaný na istý typ optického vlákna, danú vlnovú dĺžku a je charakterizovaný dvoma prenosovými parametrami:

- Vložené tlmenie (Insertion loss),
- Útlm odrazu (Return loss).

Vzdialenosť poruchy na optickom vedení môžeme určiť matematickým vzťahom (1):

$$d = \frac{c}{n} \cdot \frac{t}{v} \quad [m], \quad (1)$$

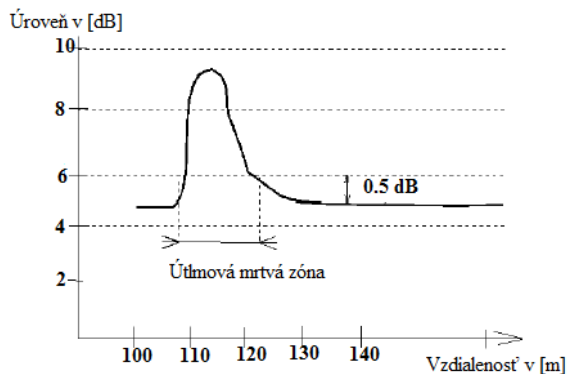
kde  $c$  je rýchlosť svetla vo vákuu ( $2,997925 \times 10^8$  m/s),  $t$  je čas od vyslania po prijatie optického impulzu a  $n$  je index lomu vlákna, ktoré je testované. OTDR využíva Rayleighov rozptyl a Fresnelov odraz na meranie stavu vlákna. Problém je, že Fresnelov odraz je desaťtisíckrát výkonovo väčší ako spätný rozptyl [2].

#### A. Rayleighov rozptyl a Fresnelov odraz

Pri meracej metóde optického spätného rozptylu hrá veľmi dôležitú úlohu Rayleighov rozptyl. Rozptylový efekt môže mať rôzny charakter, predovšetkým sa jedná o lokálnu zmenu indexu lomu, t.j. o jeho nehomogenitu. Ak sú tieto nehomogenity alebo zmeny v materiáli jadra vlákna oveľa menšie ako vlnová dĺžka žiarenia (typicky menšie ako desatina vlnovej dĺžky), dochádza na nich k Rayleighov rozptylu. Priestorové rozloženie rozptýlenej energie žiarenia a jeho celková veľkosť závisí na vlnovej dĺžke, "veľkosti" a tiež na tvare nehomogenity indexu lomu. Materiál jadra reálne vyrobeného optického vlákna nie je kryštalický, ale amorfny, čo je dané metódou jeho výroby ťahaním z preformy za tepla a následným rýchlym tuhnutím. Práve toto rýchle tuhnutie má za následok, že sa atómy nestačia usporiadať do pravidelného, priestorovo presne sa opakujúceho tvaru. Výsledkom, na molekulárnej úrovni, je neusporiadaný charakter hmoty, kedy sú v priestore mikrooblasti s väčšou hustotou a naopak niekde s menšou hustotou. Toto sa potom zodpovedajúcim spôsobom prejaví aj na fluktuácii zmeny indexu lomu, ktorá predstavuje rozptylové efekty. Rayleighov rozptyl je takmer všesmerový, optické žiarenie sa teda rovnomerne rozptyľuje vo všetkých smeroch [5].

#### B. Mŕtva zóna OTDR

Mŕtva zóna vzniká u odrazených nehomogenít na trase, Obr. 4. Spätné odrazené svetlo spôsobí saturáciu detektoru, tzv. čiastočné oslepenie, kedy detektor nie je schopný detegovať ďalšie nehomogenity na trase. Veľká mŕtva zóna vzniká v mieste pripojenia detektoru, preto je vhodné pripojiť k trase ešte predradné vlákno.



Obr. 4 Mŕtva zóna.

Odporúčanú dĺžku predradného optického vlákna môžeme vyčítať z Tab.2, kde je uvedená aj dĺžka mŕtvej zóny a dynamický rozsah vzhľadom na šírku impulzu pri prístroji OTDR 42 dB, 1310 nm.

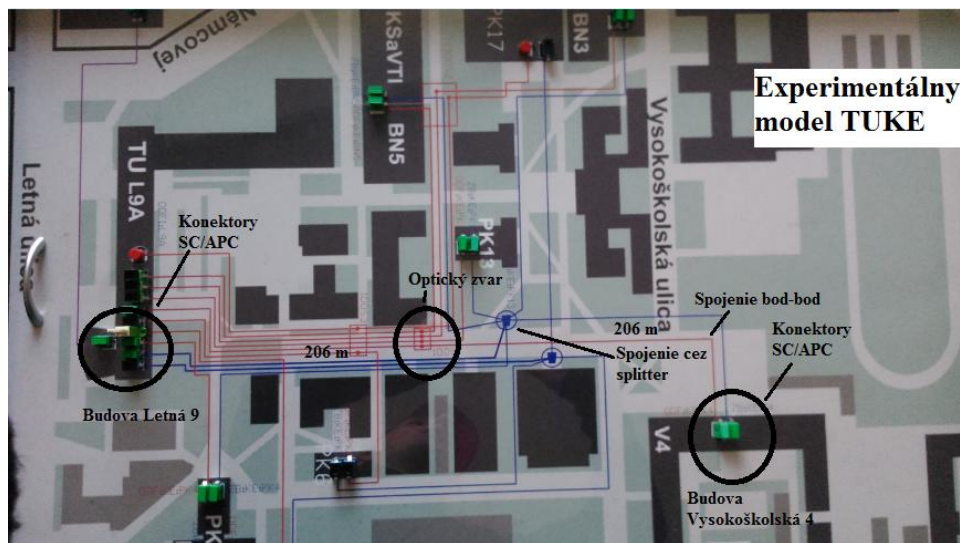
Tab. 2  
 Optické spektrum a optický výkon

Dĺžka meracieho impulzu [ns]	Dĺžka mŕtvej zóny v mieste pripojenia prístroja [m]	Dynamický rozsah [dB]	Doporučená dĺžka predradného vlákna [m]
10	30	9.7	100
30	50	12.1	200
100	70	21	
275	100	24.1	500
1000	200	27.3	
2500	400	30	1000
10000	1200	40	2000
20000	2300	49.9	3000



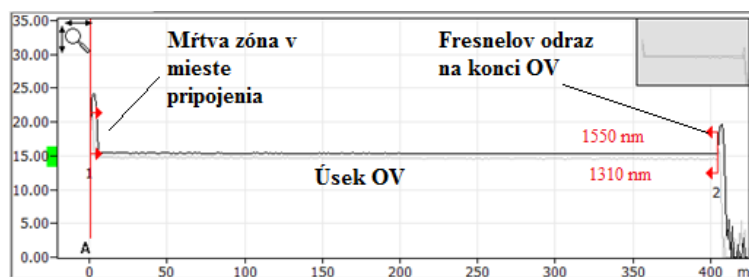
C. Namerané hodnoty OTDR

Na Obr. 5 je znázornený experimentálny model TUKE. Vo vnútri modelu sa nachádza optický 96 vlákňový kábel o dĺžke 206m. Ďalej sú tam umiestnené dve corning kazety na uchytenie ochrany zvarov.



Obr. 5 Experimentálny model TUKE.

Merania boli uskutočnené pomocou meracieho prístroja FTB-200. Namerané hodnoty sú znázornené na grafe, Obr. 6. Merania prebiehali na dvoch vlnových dĺžkach, 1310 a 1550 nm. Výsledky sú zaznamenané v Tab. 3.



Obr. 6 Graf nameraných hodnôt pre vlnové dĺžky 1310 a 1550 nm.

Tab. 3  
 Výsledky merania pre vlnové dĺžky 1310 a 1550 nm.

Číslo merania	Lok. (km)	Typ javov	Straty (dB)	Odraz (dB)	Oneskorenie (dB/km)	Kumul (dB)
<b>Meranie na 1310 nm</b>						
1	0.0000	Začiatok merania		-53.5		0.000
		Dĺžka OV (0.4046 km)	0.213		0.526	0.213
2	0.4046	Odraz		-67.7		0.213
<b>Meranie na 1550 nm</b>						
1	0.0000	Začiatok merania		-49.7		0.000
		Dĺžka OV (0.4044 km)	0.140		0.346	0.140
2	0.4044	Odraz		-58.6		0.140

Na Obr. 7 je graf nameraných hodnôt s použitím splitteru na vlnových dĺžkach 1310 nm a 1550 nm a výsledne hodnoty sú v Tab.4





Obr. 6 Graf nameraných hodnôt pre vlnové dĺžky 1310 a 1550 nm.

Tab. 3  
 Výsledky merania pre vlnové dĺžky 1310 a 1550 nm.

Číslo merania	Lok. (km)	Typ javov	Straty (dB)	Odraz (dB)	Oneskorenie (dB/km)	Kumul (dB)
<b>Meranie na 1310 nm</b>						
1	0.0000	Začiatok merania		-32.7		0.000
		Dĺžka OV (0.2068 km)	0.099		0.481	0.099
2	0.2068	Bez odrazu	9.667			9.766
		Dĺžka OV (0.2027 km)	0.182		0.896	9.948
3	0.4095	Odraz		-62.1		9.948
<b>Meranie na 1550 nm</b>						
1	0.0000	Začiatok merania		-37.0		0.000
		Dĺžka OV (0.2075 km)	0.120		0.578	0.120
2	0.2075	Bez odrazu	9.517			9.636
		Dĺžka OV (0.2056 km)	0.062		0.299	9.698
3	0.4131	Odraz				9.698

#### IV. ZÁVER

Tento článok obsahuje základné informácie o programe OptSim a o metóde meraní spätného rozptylu. Pri návrhu optickej siete sa zadávali hodnoty útlmiv z katalógu, ale pri meraní sa zistilo, že tieto hodnoty sú menšie ako udáva výrobca. Výrobcovia optických konektorov garantujú hodnotu útlmu daného konektora iba z pohľadu jeho maximálnej hodnoty. Nie je však možné zabrániť tomu, aby daný konektor mal menší alebo väčší útlm ako je jeho nominálna hodnota. Vďaka tomu sa naše namerané a odsimulované hodnoty líšia. Konektory, ktoré boli použité mali hodnotu 0,11 dB ale pri meraní OTDR sa ukázalo, že môžu mať hodnotu aj 0,08 dB. Pri použití optického rozbočovača je maximálny útlm 10,8 dB pre splittrovací pomer 1:8, ale pri meraní OTDR sa ukázal tento útlm okolo 9,5 dB. Simulácie ako aj merania sa robili na vlnových dĺžkach 1310 nm a 1550 nm.

#### POĎAKOVANIE

Táto práca bola podporovaná Kultúrnou a edukačnou grantovou agentúrou (KEGA) Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky na základe projektu č. „063TUKE-4/2013 - Použitie diaľkovo ovládaného optického vláknového refraktometra vo výučbe“ a projektu č. „006TUKE-4/2014 - Experimentálny model PON siete TUKE a jeho použitie vo výučbe“. Táto práca je tiež výsledkom zmluvy „APVV-0025-12 - Predchádzanie vplyvu stochastických mechanizmov vo vysokorýchlostných plne optických sieťach“ podporovanej Agentúrou na podporu výskumu a vývoja (APVV) Slovenskej republiky..

#### POUŽITÁ LITERATÚRA

- [1] J. Skřípský, “Měření parametrů optických konektorů ,” [online], [cit.1.5.2014]. Fakulta elektrotechniky a komunikačních technologií VUT v Brně,svazek 12,elektrorevue, ISSN 1213–1539,s.1-6.Dostupné na internete: <http://www.elektrorevue.cz/clanky/05012/index.html>
- [2] R. Hudec a kol., “Trendy vo vede a výskume informačných a komunikačných technológií,” [online], 2010, [cit.1.5.2014]. Dostupné na internete: <[http://kt.uniza.sk/populas/images/stories/upload/pub/populas\\_ebook.pdf](http://kt.uniza.sk/populas/images/stories/upload/pub/populas_ebook.pdf) 2.3.2014>.
- [3] RSoft Design Group, Inc: Manual of OptSim User Guide.
- [4] E. Ovseník, J. Turán, “Projektovanie OVKS s využitím programového balíka OptSim.” In Electrical Engineering and Informatics II. Proceeding of the Faculty of Electrical Engineering and Informatics of Technical University of Košice, Košice, Slovak Republic, March 2011, pp. 48- 53.
- [5] L. Boháč, “Měření optických vláken,” , [online], [cit.3.5.2014]. České vysoké učení technické v Praze Fakulta elektrotechnická. s.22. Dostupné nainternet: [http://data.cedupoint.cz/oppa\\_e-learning/2\\_KME/160.pdf](http://data.cedupoint.cz/oppa_e-learning/2_KME/160.pdf).

# Kombinované banky filtrov

<sup>1</sup>Ján MIHALÍK, Jozef ZAVACKÝ

<sup>1</sup> Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

jan.mihalik@tuke.sk, jozef.zavacky@tuke.sk

**Abstrakt** — Článok sa zaoberá bankami kombinovaných filtrov pre subpásmové kódovanie obrazov. V tomto článku sú navrhnuté kombinované banky filtrov zložené z banky lineárnych filtrov a banky morfológických filtrov s dokonalou rekonštrukciou, ktoré potláčajú zákmity vo vnútri obrazu a blokový efekt spôsobený blokovými procesmi realizovanými na obraze. Experimentálne výsledky boli získané pre ľubovoľné vstupné obrazy a 3-úrovňovú stromovú štruktúru, ktorá uskutočňuje pyramidový rozklad obrazu. Subobrazy z výstupu boli kvantované použitím lineárneho skalárneho kvantizátora. Presnosť rekonštrukcie obrazu pre rôzne banky filtrov sme vyhodnotili pomocou pomeru signál/šum.

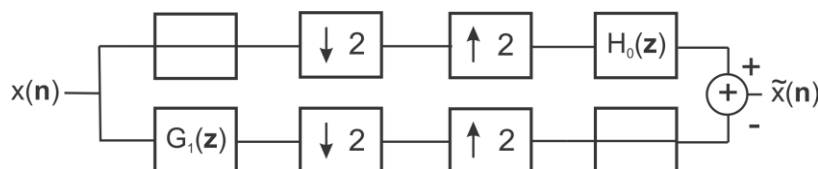
**KLúčové slová** — kombinované banky filtrov, subpásmové kódovanie, mediánový filter, transformačný zisk

## I. ÚVOD

Pri kódovaní obrazov sa môžu vyskytovať skreslenia ako napr. blokový efekt spôsobený blokovými procesmi realizovanými na obraze. Tento druh skreslenia sa neobjavuje pri použití subpásmového kódovania obrazov (SKO) [1]- [4]. Ďalším skreslením môže byť vznik zákmitov, ktoré vznikajú v dôsledku lineárnej filtrácie obrazov v ktorých sa vyskytujú oblasti s veľkým rozdielom hodnôt jasu susedných obrazových prvkov (op). V tomto článku sú uvedené také systémy analýzy a syntézy, ktorých realizáciou nevznikajú zákmity vo vnútri obrazu. Takáto analýza obrazu sa môže uskutočniť pomocou morfológického rozkladu (MR) obrazu [5]- [9]. Podmienky dokonalkej rekonštrukcie banky lineárnych filtrov (BLF) sú opísané vo frekvenčnej oblasti [10]-[12]. Keďže morfológické filtre sú nelineárne, nie je možné podmienky dokonalkej rekonštrukcie banky morfológických filtrov (BMF) opísať vo frekvenčnej oblasti. Teda podmienky dokonalkej rekonštrukcie platné pre BLF nie je možné použiť pre BMF. Použitím dolnopriepustných (DP) morfológických filtrov odstránime z obrazu malé objekty, ale hrany veľkých objektov sú zachované, čo je dôkaz prítomnosti vysokých frekvencií. Z predchádzajúceho je zjavné, že DP morfológické filtre nie je možné chápať ako filtre, ktoré z obrazu odstránia vysokofrekvenčné zložky. Môžu byť chápané ako filtre, ktoré z obrazu odstránia objekty menšie ako suport. Preto bude navrhnutá kombinovaná banka filtrov (KBF), ktorá umožňuje vzájomnú kombináciu výhod BLF a BMF.

## II. KOMBINOVANÉ BANKY FILTROV

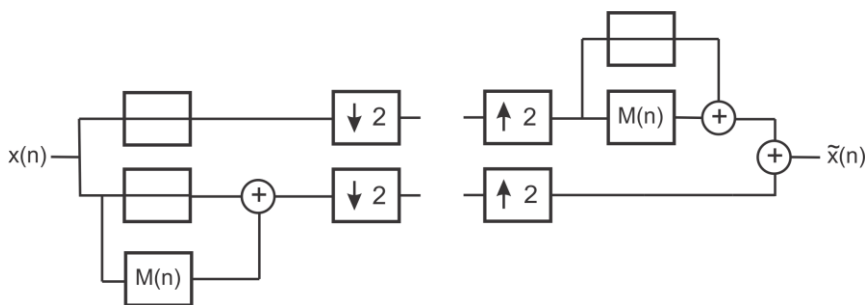
Je navrhnutá banka lineárnych filtrov (BLF) na Obr.1 a banka morfológických filtrov (BMF) na Obr.2. Tak môže byť vytvorená kombinovaná banka filtrov, ktorá je zložená z BLF a BMF.



Obr. 1 Systém analýzy a syntézy realizovaný bankou lineárnych filtrov.

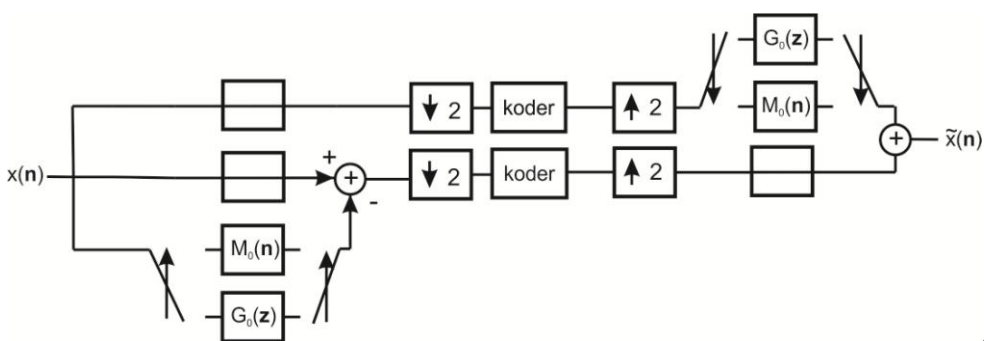
BLF má vysoký transformačný zisk (TZ) [12] pri spracovávaní textúr, ale pri spracovávaní hrán dochádza k zákmitom. Implementácia BMF je menej náročná ako implementácia BLF a nevznikajú zákmity v oblasti hrán. Preto bola navrhnutá KBF, ktorá umožňuje vzájomnú kombináciu výhod BLF a BMF. To, či v systéme analýzy a syntézy (SAS) realizovanom KBF je použitá BLF alebo BMF závisí od lokálneho charakteru obrazu. Ak obrazový prvok (op) je

súčasťou textúry, tak je použitá BLF, ak nie, tak je použitá BMF. Informácia o tom, ktorá BF bola použitá pri výpočte *op* na strane analýzy, musí byť prítomná na strane syntézy. Táto informácia má zhodné prepínanie medzi BF ako to bolo uskutočnené na strane analýzy.



Obr. 2 Systém analýzy a syntézy obrazu, banka morfológických filtrov.

Ak by informácia o spôsobe prepínania BF bola prenášaná so subobrazmi, tak by sa znížila účinnosť subpásmového kódovania obrazov (SKO). Implementácia KBF je na Obr.3

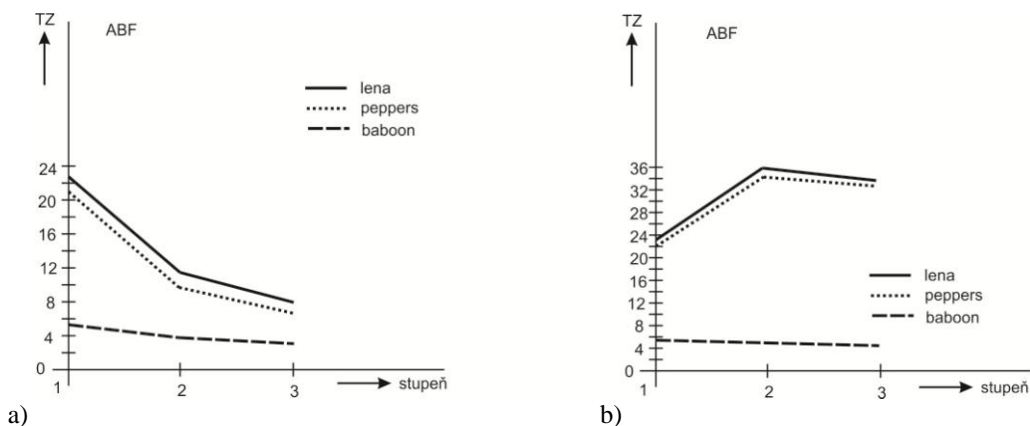


Obr. 3 Kombinovaný systém subpásmového kódovania obrazu využívajúci výhodných vlastností lineárnych a morfológických bánk filtrov.

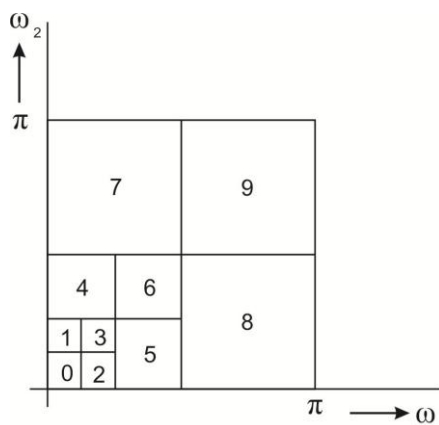
Vstupný obraz je v DP (dolnopriepustnej) vetve decimovaný s faktorom 2. V HP (hornopriepustnej) vetve je vstupný obraz filtrovaný HP lineárnym alebo morfológickým filtrom. Ak aktuálny *op* je z textúry, použije sa lineárny filter. Ak aktuálny *op* nie je z textúry, použije sa morfológický filter. Filtrovaný obraz je potom odčítaný od vstupného obrazu, čím vznikne subobraz s vysokofrekvenčným charakterom. Tento subobraz je decimovaný s faktorom 2. Do každého výstupu analýzy môže byť zaradený ďalší stupeň analýzy až kým nebude dosiahnutý požadovaný rozklad do stromovej štruktúry. Rekonštrukcia obrazu na strane syntézy je realizovaná reverzným spôsobom. V HP vetve systému syntézy sa uskutoční interpolácia subobrazu a interpolovaný subobraz je pripravený na sčítanie. V DP vetve je subobraz filtrovaný DP lineárnym alebo morfológickým filtrom. Ak aktuálny *op* je z textúry, použije sa lineárny filter. Ak *op* nie je z textúry, použije sa morfológický filter. Subobrazy z DP a HP vetvy sa sčítajú do výstupného obrazu. Tento kombinovaný systém SKO kombinuje výhodné vlastnosti bánk lineárnych a morfológických filtrov.

### III. EXPERIMENTÁLNE VÝSLEDKY

Kombinovaná BF pozostáva z BLF a BMF. Preto zvlášť budeme posudzovať vlastnosti BLF, BMF a ich kombináciou vytvorenú KBF. Najvhodnejšia BMF je Med 3x3. Je zložená z mediánových filtrov s veľkosťou suportu 3x3. K nej vyhovujúca BLF s najvyšším TZ je Hann7. Tieto banky filtrov sa preukázali ako najvhodnejšie na simuláciu KBF vytvorenej ich kombináciou. TZ tejto KBF pre vstupné obrazy Lena, peppers a baboon sú na Obr.4. Aby sa prejavili výhody KBF, je potrebné realizovať celý systém SKO. Uskutočnili sme pyramidový rozklad obrazu [1] ako je to uvedené na Obr.5.



Obr. 4 Transformačné zisky kombinovanej banky filtrov pre vstupné obrazy lena, peppers a baboon. a) Relatívny transformačný zisk b) Absolútny transformačný zisk.



Obr. 5 Pyramídový rozklad obrazu pre subpásmový kodér.

Počet pridelených bitov pre lineárny skalárny kvantizátor (LSK), ktorými kvantujeme jednotlivé subpásma je uvedený v Tabuľke.1.

Tabuľka.1  
 Počet pridelených bitov pre pásma supásmového kodéra obrazu

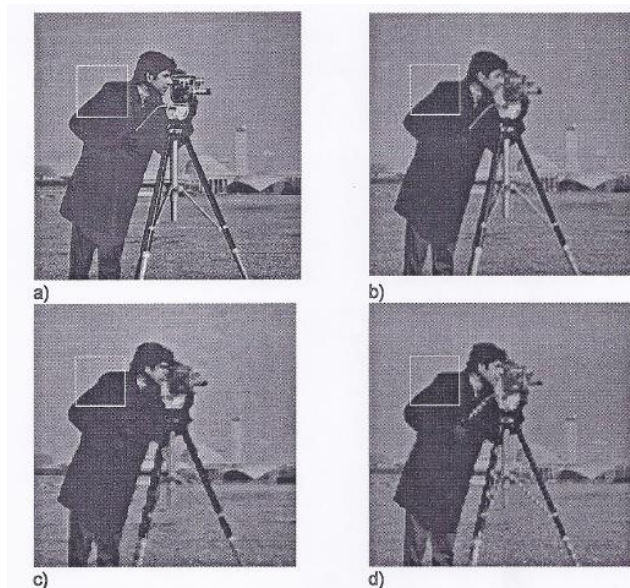
Vstupný	Počet bitov pridelených subpásmu									PBop	
obraz	0	1	2	3	4	5	6	7	8	9	[bit/OP]
lenna	4	3	3	3	0	0	0	0	0	0	0,203
peppers	4	2	2	2	0	0	0	0	0	0	0,156
baboon	4	3	4	3	3	2	3	0	0	0	0,719
camera	5	4	4	4	2	3	2	0	0	0	0,703

V Tabuľke 2 sú výsledky simulácie SKO použitím BLF, BMF a ich kombináciou vytvorenej KBF a lineárnych skalárnych kvantizátorov s prideleným počtom bitov pre subpásma v Tabuľke 1. V Tabuľke 2 sú presnosti rekonštrukcie celého obrazu, ako aj presnosti rekonštrukcie homogénnych oblastí, hrán a textúr. Z výsledkov simulácie je zrejmé, že BLF uskutočňuje najpresnejšiu rekonštrukciu obrazu, ak obraz obsahuje veľké množstvo textúr v obraze (viac ako 41% celkovej plochy obrazu). Presnosť rekonštrukcie KBF je najvyššia pre obrazy s prevahou hrán a homogénnych oblastí. Presnosť rekonštrukcie obrazu KBF závisí od percentuálneho zastúpenia textúr, hrán a homogénnych oblastí v obraze. V obraze peppers je percentuálne zastúpených 8% plochy obrazu textúrami. Teda na 8% op obrazu je aplikovaná BLF, na zvyšok sa aplikuje BMF. BMF má nižší TZ. Z tohto dôvodu rozdiel presnosti rekonštrukcie BLF a KBF je vyšší ako pri použití obrazu baboon, ktorého textúry tvoria až 63% z celkovej plochy obrazu.

Tabuľka.2  
 Výsledky simulácie subpásmového kódovania obrazu použitím kombinovanej banky filtrov

Vstupný obraz	Banka filtrov	S/Š [dB]			
		Celkovo	Homogen.	Hrany	Textúry
Lena	Hann7	21,81	27,33	19,23	15,87
	Med3x3	18,57	27,69	15,32	12,06
	KBF	21,92	28,62	19,15	15,28
peppers	Hann7	22,92	26,71	18,92	16,71
	Med3x3	19,53	26,22	13,72	12,04
	KBF	22,54	28,05	18,63	16,65
baboon	Hann7	18,51	27,86	22,75	16,31
	Med3x3	17,04	27,63	21,13	14,68
	KBF	18,35	27,65	22,15	16,18
camera	Hann7	19,41	27,59	16,17	14,7
	Med3x3	17,21	27,54	14,31	11,54
	KBF	19,14	28,16	15,52	13,82

Ak porovnáme zvlášť presnosť rekonštrukcie hrán, textúr a homogénnych oblastí, tak najvyššia presnosť rekonštrukcie pre všetky vstupné obrazy je dosiahnutá pre homogénne oblasti. Ak by obraz obsahoval len homogénne oblasti, tak pri rozklade obrazu pomocou systému analýzy sa celá energia sústreďuje do malej frekvenčnej plochy v transformovanom priestore. Prejaví sa to vysokým TZ a teda s vysokou účinnosťou kódovania. Naopak textúry sú charakteristické prítomnosťou vo viacerých pásmach, čo znižuje TZ a presnosť rekonštrukcie pre konštantnú prenosovú rýchlosť. Výhodou KBF je použitie BLF s vyšším TZ pri spracovaní textúr. Hrany sú filtrované BMF, čo odstraňuje zákmitý vo výstupnom obraze. Na obr.6 sú výstupné obrazy pri použití BLF, BMF a KBF a vstupného obrazu kamera.

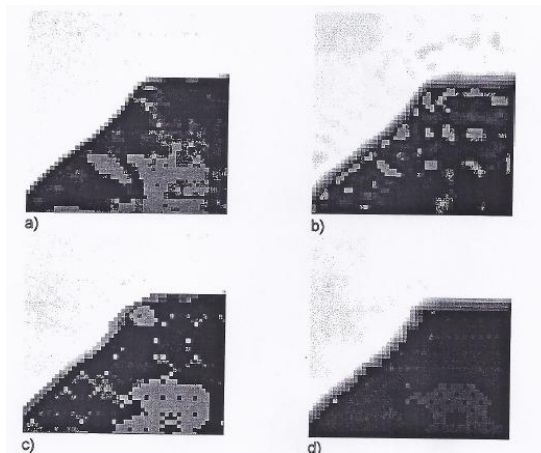


Obr. 6 Výstupné obrazy simulácie subpásmového kódovania obrazu pre rôzne banky filtrov s označením skúmaného detailu a) vstupný obraz b) banka lineárnych filtrov c) banka morfológických filtrov d) kombinovaná banka.

Vo výstupných obrazoch sú rámčeky, ktoré označujú detaily zobrazené na Obr.7. Z Obr.6 a Obr.7 vyplýva, že BLF spôsobuje zákmitý vo výstupnom obraze, ktoré pôsobia rušivo.



Implementáciou BMF nevznikajú zákmity, ale je dosiahnutá nižšia presnosť rekonštrukcie obrazu (viď. Tabuľka 2). Nižšia presnosť rekonštrukcie je spôsobená a tým, že použitý morfológický filter (mediánový filter) mení pôvodnú veľkosť a tvar objektov v obraze. Zreteľne je to na obrysoch objektov, ktoré sú pod iným uhlom ako násobok  $\pi/4$ . Je to evidentné na stojane kamery Obr.6c. Ďalším nedostatkom BMF je nedostatočná rekonštrukcia textúr (tráva, kamera).



Obr. 7 Detaily vstupných obrazov simulácie subpásmového kódovania obrazu pre rôzne banky filtrov a) vstupný obraz b) banka lineárnych filtrov c) banka morfológických filtrov d) kombinovaná banka filtrov.

KBF podľa očakávania lepšie rekonštruje textúry v obraze kamera v porovnaní s BMF. Presnosťou rekonštrukcie textúr sa približia BLF. Použitím KBF sa opäť prejavuje nedostatok morfológických filtrov a to dodržiavame tvaru a veľkosti objektov. Podobne ako pri BMF sa to najviac prejavuje pri hranách s uhlom sklonu iným ako násobok  $\pi/4$ , Obr.6d. Detailné obrazy na Obr.7 podrobne dokumentujú a potvrdzujú závery zo simulácie SKO s použitím BLF, BMF a KBF.

#### IV. ZÁVER

Získané experimentálne výsledky potvrdili výhodné vlastnosti kombinovaných filtrov pri subpásmovom kódovaní obrazov. Z výsledkov simulácie je zrejmé, že BLF uskutočňuje najpresnejšiu rekonštrukciu obrazu, ak obraz obsahuje veľké množstvo textúr v obraze (viac ako 41% celkovej plochy obrazu). Presnosť rekonštrukcie KBF je najvyššia pre obrazy s prevahou hrán a homogénnych oblastí. Presnosť rekonštrukcie obrazu KBF závisí od percentuálneho zastúpenia textúr, hrán a homogénnych oblastí v obraze. Výhodou KBF je použitie BLF s vyšším TZ pri spracovaní textúr. Hrany sú filtrované BMF, čo odstraňuje zákmity vo výstupnom obraze. Použitím KBF sa opäť prejavuje nedostatok morfológických filtrov a to dodržiavame tvaru a veľkosti objektov.

#### LITERATÚRA

- [1] J. Mihalík, Kódovanie obrazu vo videokomunikáciách. Mercury-Smekal, Košice, 2001. ISBN 80-89061-47-8.
- [2] G. Jovanovic-Dolecek, *Multirate Systems: Design and Applications*, Hershey PA: Idea Group Publishing, 2002.
- [3] J. Mihalík, – J. Zavacký – J. Dzivý, J, Perfect Reconstruction 2D QMF Bank for Subband Image Coding, *Journal of Electrical Engineering*, Vol.47, No.9, 1996, p.195-201.
- [4] T. Acharya – A. K. Ray, *Image Processing: Principles and Applications*, John Wiley & Sons, 2005.
- [5] O. Egger – M. Kunt, High Compression Image Coding Using an Adaptive Morphological Subband Decomposition. *Proceedings of the IEEE*, Vol. 83, No. 2, 1995, pp.272-278.
- [6] J. R. Casas – L. Torres, A region-based subband coding. *Signal processing: Image communication*, Vol. 10, July 1997, pp. 173-200.
- [7] Ch. S. Pei – F.Ch. Chen, Hierarchical image representation by mathematical morphology subband decomposition. *Pattern recognition letters*, 1995, pp. 183-192.
- [8] P. Salembier, Morphological multiscale segmentation for image coding. *EURASIP Signal Processing*, Vol. 38, No. 3, 1994, pp. 359-386.
- [9] S. Serveto – K. Ramchardran – M. Orchard, Image coding based on a morphological representation of wavelet data. *IEEE Trans. On Image Processing*, Vol.8, No. 9, 1999, pp. 1161-1174.
- [10] J. Mihalík – J. Zavacký, Diskrétné spracovanie signálov. LČSOV FEI TU Košice, 2011, ISBN 978-80-553-0730-5, (319 strán).
- [11] J. Mihalík – J. Zavacký – I. Gladišová, Mnohokanálové diskrétné systémy. *Slaboproudý obzor*, roč. 65, č.4, 2009, str.P1-P8, ISSN 0037-668
- [12] J. Zavacký – J. Mihalík, Banky filtrov. LČSOV FEI TU Košice, 2012, ISBN 978-80-553-0938-5, (93 s.).

# Konverzia časovej lifting implementácie diskkrétnej waveletovej transformácie do z-roviny

<sup>1</sup>Jozef ZAVACKÝ, Ján MIHALÍK

<sup>1</sup> Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

jan.mihalik@tuke.sk, jozef.zavacky@tuke.sk

**Abstrakt** — Článok sa zaoberá konverziou časovej lifting implementácie DWT pomocou dvojkanálového diskkrétneho systému s dokonalou rekonštrukciou (2KDS DR) do oblasti Z-transformácie. Najprv sme teoreticky rozpracovali všeobecný postup uvedenej konverzie kedy sa na základe 2KDS DR určí polyfázová matica Laurantových polynómov premenných  $z$  a  $z^{-1}$  a následne sa vykoná jej faktorizácia na matice s rozmerom  $2 \times 2$ . Pomocou nich dostaneme transformovanú lifting implementáciu (LI) reprezentovanú blokovou schémou na základe ktorej možno odvodiť rovnice pre výpočet detailových a aproximačných koeficientov. Uvedený všeobecný postup sme nakoniec aplikovali na transformovanú LI DWT s biortogonálnou bankou filtrov (5,3) a (9,7).

**Kľúčové slová** — konvolučná metóda, lifting implementácia, faktorizácia, banka filtrov

## I. ÚVOD

Lifting metóda [1]-[2] implementácie diskkrétnej waveletovej transformácie (DWT) v porovnaní s klasickou konvolučnou metódou [3]-[5] má nasledujúce výhody. Urýchluje výpočet, preto je tiež označovaná ako rýchla lifting DWT. Dáva možnosť vykonať výpočty bez použitia prídavnej pamäte, to znamená, že pôvodný diskrétny signál môže byť priamo nahradený jeho DWT. Spätná DWT môže byť získaná bezprostredne z priamej DWT, zámenou poradia operácií a výmenou každého znamienka plus za mínus a naopak. Umožňuje ľahko získať celočíselnú DWT, ktorá má jednoduchú hardverovú implementáciu a umožňuje bezstratovú kompresiu. Lifting implementácia DWT začína s triviálnym waveletom a postupne zostavuje nový wavelet so zlepšenými vlastnosťami pridávaním nových básových funkcií. To bolo zdrojom inšpirácie pre názov „lifting implementácia“, čo by sme mohli nazvať implementácia postupného vylepšovania vlastností nového waveletu. Ak uvažujeme postupnosť  $c_j(n)$  s  $2^j$  vzorkami (rozkladovými koeficientami), túto môžeme transformovať na postupnosti hrubšej aproximácie  $c_{j-1}(n)$  a detailov  $d_{j-1}(n)$ . Jeden stupeň časovej lifting implementácie priamej DWT pozostáva obvykle z troch krokov: rozdelenia, predikcie a korekcie. Vo všeobecnosti operator korekcie ako aj predikcie závisia od špecifikácie konkrétnej LI priamej DWT. Ďalej bude ukázaný spôsob konverzie časovej LI DWT pomocou dvojkanálového diskkrétneho systému s dokonalou rekonštrukciou do oblasti Z-transformácie.

## II. IMPLEMENTÁCIA DWT LIFTING METÓDOU V OBLASTI Z TRANSFORMÁCIE

Transformovaná LI DWT sa získa Z-transformáciou jej časovej LI. Vo všeobecnosti vstupná postupnosť  $c(n)$  jedného stupňa časovej LI priamej DWT, nezávisle od  $j$ , sa na začiatku rozdelí na párnú  $c^0(n)$  a nepárnú  $d^0(n)$  postupnosť. Toto pomocou Z-transformácie sa vykoná takto [6]

$$C(z) = C^0(z^2) + z^{-1}D^0(z^2) \quad (1)$$

kde

$$C^0(z) = \sum_n c(2n) z^{-n} \quad (2)$$

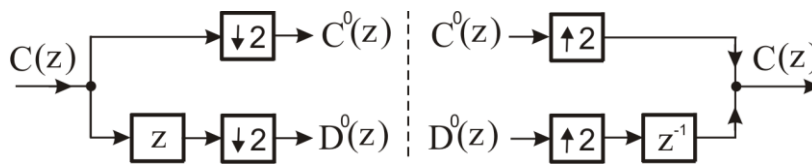
$$D^0(z) = \sum_n c(2n+1) z^{-n} \quad (3)$$

Vidíme, že  $C^0(z)$  dostaneme z pôvodného signálu decimáciou s faktorom dva.  $D^0(z)$  dostaneme z daného signálu  $c(n)$  najprv jeho posunutím o jednu periódu  $T_d$  doľava, a potom decimáciou opäť s faktorom decimácie dva. Podobne  $D^0(z)$  bude Z-transformácia decimovaného výstupného signálu s tým istým faktorom, ale až po jeho posunutí o jednu periódu doľava. Na základe vzťahu (2) a vlastnosti Z- transformácie o posunutí postupnosti máme

$$C^0(z) = \frac{1}{2} \left( C(z^{1/2}) + C(-z^{1/2}) \right) \quad (4)$$

$$D^0(z) = \frac{z^{1/2}}{2} \left( C(z^{1/2}) - C(-z^{1/2}) \right) \quad (5)$$

Toto rozdelenie vstupnej postupnosti na párnú a nepárnu postupnosť v z-rovine môžeme zobrazit' tak ako to vidno na Obr.1a. Naopak ich zlúčenie za účelom jej spätnej obnovy v z-rovine možno vidieť na Obr.1b . Výsledkom tohto zlúčenia je  $C(z)$ , ktoré získame z  $C^0(z)$  a  $D^0(z)$  najprv pomocou interpolácie s faktorom 2, potom posunutím o jednu periódu doprava pre zložku  $D^0(z^2)$  (násobenie s  $z^{-1}$ ) a nakoniec sčítaním takto získaných zložiek ako to ilustruje Obr.1b.



Obr.1 a) Rozdelenie vstupnej postupnosti na párnú a nepárnu postupnosť,  
 b) ich opätovné zlúčenie za účelom jej spätnej obnovy v z-rovine.

Ďalej v druhom kroku implementujeme operáciu predikcie spolu s odčítaním, ako to vyplýva z jedného stupňa časovej LI priamej DWT vstupnej postupnosti  $c_j(n)$  [2]. Výsledkom týchto operácií v z-rovine bude  $D^0(z) - P_1(z) \cdot C^0(z)$ , kde  $P_1(z) = p_1(1+z)$ . Najprv vypočítame súčin

$$\begin{aligned} P_1(z)C^0(z) &= p_1(1+z) \sum_n c^0(n)z^{-n} = p_1 \sum_n c^0(n)z^{-n} + p_1 \sum_n c^0(n)z^{-n+1} \\ &= \sum_n p_1 \left( c^0(n) + c^0(n+1) \right) z^{-n} \end{aligned} \quad (6)$$

a potom rozdiel

$$D^0(z) - P_1(z)C^0(z) = \sum_n \left( d^0(n) - p_1(c^0(n) + c^0(n+1)) \right) z^{-n} \quad (7)$$

z ktorého vyplýva, že rov.(7) predstavuje Z-transformáciu  $D^1(z)$  z  $d^1(n)$ , pre prvú iteráciu. Prechod od  $(C^0(z), D^0(z))$  k  $(C^0(z), D^0(z) - P_1(z)C^0(z))$  môžeme vyjadriť v maticovom tvare nasledovne

$$\begin{bmatrix} C^0(z) \\ D^1(z) \end{bmatrix} = \begin{bmatrix} C^0(z) \\ D^0(z) - P_1(z)C^0(z) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -P_1(z) & 1 \end{bmatrix} \begin{bmatrix} C^0(z) \\ D^0(z) \end{bmatrix} \quad (8)$$

Nakoniec tretí krok korekcie spolu so sčítaním dáva v z-rovine výsledok  $C^0(z) + K_1(z)D^1(z)$ , kde  $K_1(z) = k_1(1+z^{-1})$ . Tento výsledok predstavuje Z-transformáciu  $C^1(z)$  z  $c^1(n)$ , pre prvú iteráciu tohto kroku. Potom prechod od  $(C^0(z), D^1(z))$  k  $(C^0(z) + K_1(z)D^1(z), D^1(z))$  v maticovom tvare bude

$$\begin{bmatrix} C^1(z) \\ D^1(z) \end{bmatrix} = \begin{bmatrix} C^0(z) + K_1(z)D^1(z) \\ D^1(z) \end{bmatrix} = \begin{bmatrix} 1 & K_1(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} C^0(z) \\ D^1(z) \end{bmatrix} \quad (9)$$

Nakoniec po vykonaní všetkých iterácií sa prevedie normovanie v závislosti od druhu BF pomocou konštánt  $A_1 > 0, A_2 > 0$ , ktoré zodpovedá násobeniu maticou

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \quad (10)$$



Vo všeobecnosti rovnaký postup implementácie jednotlivých krokov príslušného stupňa transformovanej LI možno použiť pre ľubovoľnú iteráciu. Potom  $i$ -tej iterácii predikčného kroku zodpovedá v  $z$ -rovine násobenie maticou

$$\mathbf{L}_i(z) = \begin{bmatrix} 1 & 0 \\ -P_i(z) & 1 \end{bmatrix} \quad (11)$$

a podobne  $i$ -tej iterácii korekčného kroku násobenie maticou

$$\mathbf{M}_i(z) = \begin{bmatrix} 1 & K_i(z) \\ 0 & 1 \end{bmatrix} \quad (12)$$

kde  $P_i(z)$  a  $K_i(z)$  sú Laurentove polynómy premenných  $z$  a  $z^{-1}$  [6]. Takýto postup po  $N$  iteráciách, resp.  $N$  násobeniach maticami  $\mathbf{L}_i(z)$  a  $\mathbf{M}_i(z)$  dáva výslednú maticu

$$\mathbf{U}(z) = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} 1 & K_N(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -P_N(z) & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & K_1(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -P_1(z) & 1 \end{bmatrix} \quad (13)$$

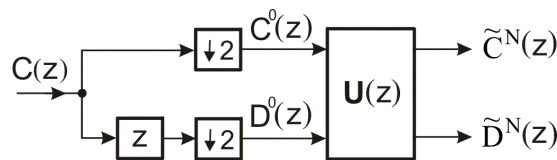
Poradie jednotlivých matic je určené poradím v akom sa vykonávajú jednotlivé kroky a ich iterácie. Najprv krok predikcie, potom krok korekcie, opakovane  $N$  krát. Po inverzii matice  $\mathbf{U}(z)$  z rov.(13) dostaneme jej inverznú maticu

$$\mathbf{V}(z) = \begin{bmatrix} 1 & 0 \\ P_1(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & -K_1(z) \\ 0 & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ P_N(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & -K_N(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_2 & 0 \\ 0 & A_1 \end{bmatrix} \quad (16)$$

Vynásobením všetkých matic v rov. (16) získame maticu, ktorej prvky sú Laurentove polynómy a v štandardnej forme má tvar

$$\mathbf{U}(z) = \begin{bmatrix} U_{00}(z) & U_{01}(z) \\ U_{10}(z) & U_{11}(z) \end{bmatrix} \quad (17)$$

Z tohto tvaru však nie je tak zrejماً jej invertibilita ako z tvaru v rov. (16). Preto ako pre bežné matice, môžeme použiť jej determinant pre rozhodnutie, či je invertovateľná. Potom bude invertovateľná vtedy a len vtedy, ak jej determinant je monomial. Výsledné zapojenie jedného stupňa transformovanej LI priamej DWT je na Obr. 2.

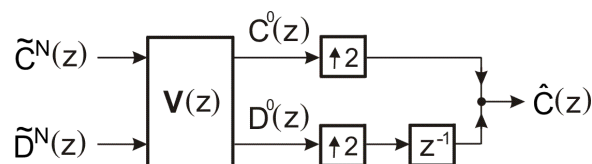


Obr.2 Zapojenie jedného stupňa transformovanej LI priamej DWT.

Na základe Obr. 2 platí maticová rovnica

$$\begin{bmatrix} \tilde{C}^N(z) \\ \tilde{D}^N(z) \end{bmatrix} = \begin{bmatrix} U_{00}(z) & U_{01}(z) \\ U_{10}(z) & U_{11}(z) \end{bmatrix} \begin{bmatrix} C^0(z) \\ D^0(z) \end{bmatrix} \quad (18)$$

ktorá predstavuje polyfázovú reprezentáciu jedného stupňa transformovanej LI priamej DWT. Analogickým postupom možno získať zapojenie jedného stupňa transformovanej LI spätnej DWT, uvedené na Obr.3. Toto zapojenie zodpovedá obrátenému poradiu operácií predikcie a korekcie so súčasnou zámienou znamienok. Vtedy matica  $\mathbf{V}(z)$  je daná inverznou maticou z  $\mathbf{U}(z)$ , tak ako je to vidno z rov.(16). Nakoniec zlúčenie párných a nepárných zložiek sa vykoná v tomto obrázku, tak ako je to samostatne ukázané na Obr. 1b.

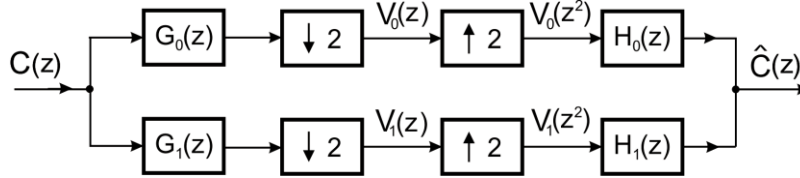


Obr.3 Zapojenie jedného stupňa transformovanej LI spätnej DWT.

Z Obr.3 polyfázová reprezentácia jedného stupňa transformovanej LI spätnej DWT bude

$$\begin{bmatrix} C^0(z) \\ D^0(z) \end{bmatrix} = \begin{bmatrix} V_{00}(z) & V_{01}(z) \\ V_{10}(z) & V_{11}(z) \end{bmatrix} \begin{bmatrix} \tilde{C}^N(z) \\ \tilde{D}^N(z) \end{bmatrix} \quad (19)$$

Kaskádnym zapojením schém z Obr. 2 a 3 dostaneme 2KDS s DR ( $C(z) = \hat{C}(z)$ ), ktorej ekvivalentné zapojenie je na Obr. 4 a predstavuje konvolučnú implementáciu dvojkanálovej DWT.



Obr. 4 Zapojenie 2KDS s DR pomocou banky filtrov analýzy a syntézy.

Za predpokladu, že obidve zapojenia sú ekvivalentné dostaneme vzťah, resp. spôsob prepočtu transformovanej LI na klasickú konvolučnú implementáciu DWT. V súvislosti s tým pre Z-transformácie výstupných signálov decimátorov na Obr. 4 platí

$$V_0(z) = \frac{1}{2} \left( G_0(z^{1/2})C(z^{1/2}) + G_0(-z^{1/2})C(-z^{1/2}) \right) \quad (20)$$

$$V_1(z) = \frac{1}{2} \left( G_1(z^{1/2})C(z^{1/2}) + G_1(-z^{1/2})C(-z^{1/2}) \right) \quad (21)$$

Podobne Z-transformácie ekvivalentných signálov na Obr. 2 vypočítame z rov.(18) a pre ne po substitúcii z rov. (4) a (5) máme

$$\begin{bmatrix} \tilde{C}^N(z) \\ \tilde{D}^N(z) \end{bmatrix} = \mathbf{U}(z) \begin{bmatrix} \frac{1}{2} \left( C(z^{1/2}) + C(-z^{1/2}) \right) \\ \frac{z^{1/2}}{2} \left( C(z^{1/2}) - C(-z^{1/2}) \right) \end{bmatrix} \quad (22)$$

Pretože ľavé strany predchádzajúcich rovnice sú rovnaké musia sa rovnať aj ich pravé strany z čoho po substitúcii kvadrátom  $z^2$  bude  $V_0(z^2) = \tilde{C}^N(z^2)$  a potom máme

$$\begin{aligned} V_0(z^2) &= \frac{1}{2} (G_0(z)C(z) + G_0(-z)C(-z)) = \\ \tilde{C}^N(z^2) &= U_{00}(z^2) \frac{1}{2} (C(z) + C(-z)) + U_{01}(z^2) \frac{z}{2} (C(z) - C(-z)) \end{aligned} \quad (23)$$

čo vedie k vzťahu

$$G_0(z) = U_{00}(z^2) + zU_{01}(z^2) \quad (24)$$

Z tohto vzťahu vidno, že pred výrazom  $U_{01}(z^2)$  sme dostali  $z$  a nie  $z^{-1}$ . Vzťah pre  $G_0(z)$ ,  $G_1(z)$  dostaneme analogickým spôsobom, kedy  $V_1(z^2) = \tilde{D}^N(z^2)$  z čoho dostaneme

$$G_1(z) = U_{10}(z^2) + zU_{11}(z^2) \quad (25)$$

Vzťahy pre  $H_0(z)$  a  $H_1(z)$  možno získať na základe podmienky dokonalej rekonštrukcie, kedy  $\hat{C}(z) = C(z)$  pre ktorú platí

$$\begin{aligned} H_0(z)G_0(z) + H_1(z)G_1(z) &= 2 \\ H_0(z)G_0(-z) + H_1(z)G_1(-z) &= 0 \end{aligned} \quad (26)$$

Výsledné vzťahy potom budú

$$\begin{aligned} H_0(z) &= V_{00}(z^2) + z^{-1}V_{10}(z^2) \\ H_1(z) &= V_{01}(z^2) + z^{-1}V_{11}(z^2) \end{aligned} \quad (27)$$

Požiadavka dokonalej rekonštrukcie pre polyfázové reprezentácie jedného stupňa transformovanej LI priamej a spätnej DWT je daná výpočtom  $\mathbf{V}(z) = \mathbf{U}^{-1}(z)$ . Naopak z konvolučnej implementácie DWT pomocou 2KDS s DR možno nájsť jej transformovanú LI. Vtedy pomocou rov. (24), (25), (27) a vzťahu medzi maticami  $\mathbf{U}(z)$ ,  $\mathbf{V}(z)$  vypočítame maticu  $\mathbf{U}(z)$ . Následne túto maticu faktorizujeme a normujeme podľa rov. (13), aby sme na základe nej mohli získať transformovanú LI DWT. Vo všeobecnosti determinant matice v tvare danom rov. (17), t.j.  $\det \mathbf{U}(z) = cz^{2k+1}$ , zatiaľ čo v rov. (13) je tento determinant rovný 1. Pomocou operácií normovania a posúvania v čase o nepárnu hodnotu možno dosiahnuť, aby tento determinant bol vždy rovný jednej a až potom sa vykonáva faktorizácia matice  $\mathbf{U}(z)$ . Za predpokladu, že  $\mathbf{U}(z)$  je matica Laurentových polynómov, normovaná tak, že  $\det \mathbf{U}(z) = 1$ . Potom existuje konštanta  $A \neq 0$  ( $A_1=A$ ,  $A_2=1/A$ ) a Laurantove polynómy  $P_1(z)$ , ... ,  $P_N(z)$ ,  $K_1(z)$ , ...  $K_N(z)$  také, že túto maticu možno vyjadriť v tvare rov. (13).

Ďalej budeme aplikovať vyššie uvedené výsledky na transformovanú LI DWT s BF(5,3) a BF(9,7). Uvažujme najprv BF(5,3) [2] pre ktorú prenosové funkcie filtrov analýzy sú

$$G_0(z) = -\frac{1}{8}z^2 + \frac{1}{4}z + \frac{3}{4} + \frac{1}{4}z^{-1} + \frac{1}{8}z^{-2} \quad (28)$$

$$G_1(z) = -\frac{1}{2}z^2 + z - \frac{1}{2} \quad (29)$$

Tieto rovnice upravíme do polyfázového tvaru

$$G_0(z) = -\frac{1}{8}z^2 + \frac{3}{4} - \frac{1}{8}z^{-2} + z\left(\frac{1}{4} + \frac{1}{4}z^{-2}\right) = U_{00}(z^2) + zU_{01}(z^2) \quad (30)$$

$$G_1(z) = -\frac{1}{2}z^2 - \frac{1}{2} + z = U_{10}(z^2) + zU_{11}(z^2) \quad (31)$$

Potom matica

$$\mathbf{U}(z) = \begin{bmatrix} -\frac{1}{8}z + \frac{3}{4} - \frac{1}{8}z^{-1} & \frac{1}{4} + \frac{1}{4}z^{-1} \\ -\frac{1}{2} - \frac{1}{2}z & 1 \end{bmatrix} \quad (32)$$

ktorej  $\det \mathbf{U}(z) = 1$ . Faktorizáciou  $\mathbf{U}(z)$  dostaneme

$$\mathbf{U}(z) = \begin{bmatrix} 1 & \frac{1}{4}(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1+z) & 1 \end{bmatrix} = \begin{bmatrix} 1 & K_1(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -P_1(z) & 1 \end{bmatrix} \quad (33)$$

Po substitúcii takto faktorizovanej normovanej matice  $\mathbf{U}(z)$  do rov. (18) pre jednu iteráciu, kedy  $N=1$  máme

$$\begin{bmatrix} \tilde{C}^1(z) \\ \tilde{D}^1(z) \end{bmatrix} = \begin{bmatrix} 1 & K_1(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -P_1(z) & 1 \end{bmatrix} \begin{bmatrix} C^0(z) \\ D^0(z) \end{bmatrix} \quad (34)$$

Teraz postupne vykonáme maticové násobenie v rov. (34) a po prvom vynásobení dostaneme

$$\begin{bmatrix} 1 & 0 \\ -P_1(z) & 1 \end{bmatrix} \begin{bmatrix} C^0(z) \\ D^0(z) \end{bmatrix} = \begin{bmatrix} C^0(z) \\ D^0(z) - C^0(z)P_1(z) \end{bmatrix} = \begin{bmatrix} C^0(z) \\ D^1(z) \end{bmatrix} \quad (35)$$

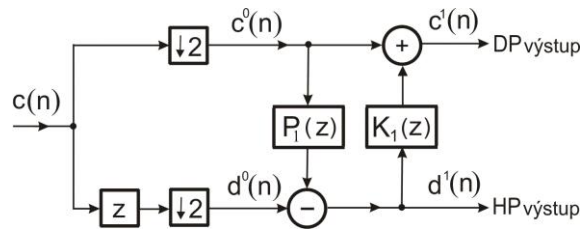
kde  $D^1(z) = D^0(z) - P_1(z)C^0(z) = D^0(z) - C^0(z)\frac{1}{2}(1+z)$  (36)

Po spätnej Z-transformácii rov.(36) dostaneme prvú rovnice v časovej oblasti

$$d^1(n) = d^0(n) - \frac{1}{2}[c^0(n) + c^0(n+1)] \quad (37)$$

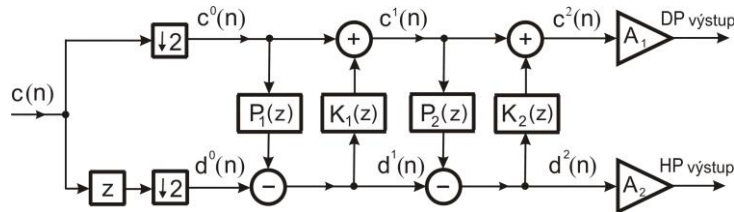
$$c^1(n) = c^0(n) + \frac{1}{4}[d^1(n-1) + d^1(n)] \quad (38)$$

Výsledná transformovaná LI jedného stupňa priamej DWT s BF(5,3) je na Obr. 5, kde  $P_1(z) = 0,5(1+z)$  a  $K_1(z) = -0,25(1+z^{-1})$ .



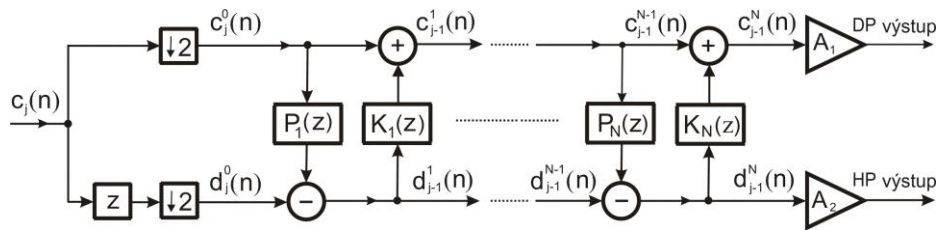
Obr.5 Transformovaná LI jedného stupňa priamej DWT s BF(5,3).

Analogickým postupom pre BF (9,7) [2] možno získať transformovanú LI jedného stupňa priamej DWT na Obr. 6, kde  $P_1(z)=-p_1(1+z)$ ,  $K_1(z)=k_1(1+z^{-1})$ ,  $P_2(z)=-p_2(1+z)$ ,  $K_2(z)=k_2(1+z^{-1})$ .

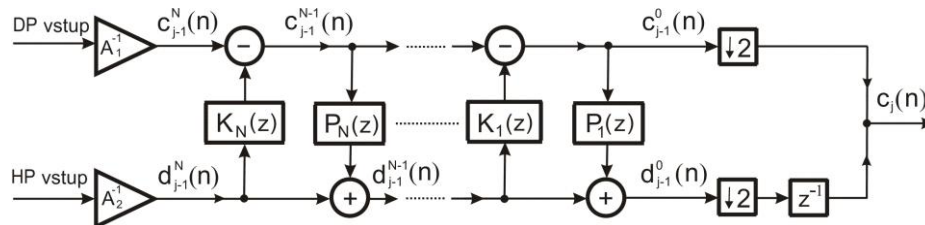


Obr.6 Transformovaná LI jedného stupňa priamej DWT s BF(9,7).

Zovšeobecním pre N iterácií a j-tý stupeň transformovanej LI DWT dostaneme zapojenie na Obr. 7 pre priamu a na Obr. 8 pre spätnú DWT.



Obr. 7 Zovšeobecnená transformovaná LI j-tého stupňa priamej DWT.



Obr. 8 Zovšeobecnená transformovaná LI j-tého stupňa spätnej DWT.

### III. ZÁVER

Z konvolučnej implementácie DWT pomocou 2KDS s DR možno nájsť jej transformovanú LI. Výsledok faktorizácie nie je jediný možný a v literatúre možno nájsť pre danú DWT celkom odlišné výsledky faktorizácie. Záleží na aplikácii, kedy si môžeme zvoliť faktorizáciu s najmenšou hodnotou N, alebo s konštantami  $A_1$ ,  $A_2$  blízskymi jednej, alebo zachovať symetriu. Výhoda transformovanej LI DWT v porovnaní s jej klasickou konvolučnou implementáciou pomocou 2KDS s DR je tá, že je možné pomerne ľahko nájsť filtre analýzy  $G_0(z)$ ,  $G_1(z)$  a syntézy  $H_0(z)$ ,  $H_0(z)$ .

### LITERATÚRA

- [1] W. Sweldens, The Lifting Scheme: A Construction of Second Generation Wavelets. *Siam J. Math. Anal.*, 1997, Vol. 29, No. 2, pp. 511-546.
- [2] J. Zavacký - J. Mihalík, - I. Gladišová, Implementácia diskretnej waveletovej transformácie v štandarde JPEG-2000. *Slaboproudý obzor*, roč. 63, č.3-4, 2007, s.5-9, ISSN 0037-668X
- [3] J. Mihalík - J. Zavacký - I. Gladišová, Mnohokanálové diskretné systémy. *Slaboproudý obzor*, roč. 65, č.4, 2009, str.P1-P8, ISSN 0037-668X.
- [4] J. Zavacký - J. Mihalík, Banky filtrov. LČSOV FEI TU Košice, 2012, ISBN 978-80-553-0938-5, (93 s.).
- [5] J. Mihalík - J. Zavacký - J. Dzivý, Perfect Reconstruction 2D QMF Bank for Subband Image Coding, *Journal of Electrical Engineering*, Vol.47, No.9, 1996, pp.195-201.
- [6] A. Jensen - A. la Cour- Harbo, Ripples in Mathematics, The Discrete Wavelet Transform. Springer-Verlag, Berlin Heidelberg, 2000, ISBN 3-540-41662-5

# Vizualizačná platforma VizAlgo a jej rozšírenie

<sup>1</sup>Lukáš KUKLA, <sup>2</sup>Slavomír ŠIMOŇÁK

Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovenská Republika

<sup>1</sup>lukas.kukla@student.tuke.sk, <sup>2</sup>slavomir.simonak@tuke.sk

**Abstrakt** — Obsahom tejto práce je rozšírenie prostredia na vizualizáciu činnosti algoritmov VizAlgo, ktoré umožňuje grafickým spôsobom demonštrovať činnosť daného algoritmu, prípadne údajovej štruktúry. Prvoradými cieľmi tejto práce je tvorba nových zásuvných modulov, prídanie novej formy zobrazenia vizualizácie triediacich algoritmov a podpory pre vizualizáciu nových druhov údajových štruktúr. Pri riešení uvedených cieľov bolo potrebné vytvoriť nové triedy implementujúce dané algoritmy a upraviť jadro pôvodnej aplikácie pre prídanie podpory nového zobrazenia vizualizácie a vizualizácie nových typov algoritmov alebo údajových štruktúr. To umožnilo rozšíriť využiteľnosť aplikácie pri výučbe algoritmov a údajových štruktúr.

**Kľúčové slová** — algoritmus, VizAlgo, vizualizácia, údajová štruktúra, zásuvný modul

## I. ÚVOD

Väčšina algoritmov a údajových štruktúr, ktoré sú používané v informatike sa môže zdať na prvý pohľad ťažšie pochopiteľná. Na ich ľahšie pochopenie nám teda môže poslúžiť grafická vizualizácia ich činnosti. Táto práca sa zaoberá už existujúcimi softvérovými riešeniami z oblasti vizualizácie, vizualizačnou platformou VizAlgo a jej rozšírením. Primárnym cieľom práce bola analýza súčasného stavu aplikácie VizAlgo a následne jej rozšírenie a to nie len prostredníctvom nových zásuvných modulov ale aj o podporu nových foriem vizualizácie a nových údajových štruktúr. Keďže aplikácia vo svojom základe umožňovala vizualizáciu triediacich algoritmov a abstraktného dátového typu zásobník, jej rozšírením bolo umožnené používateľovi pochopenie činnosti nových algoritmov a údajových štruktúr.

## II. PREHEAD EXISTUJÚCICH RIEŠENÍ

Pri analýze existujúcich vizualizačných systémov som narazil na mnoho odvetví využitia vizualizácie ako je napríklad vizualizácia grafových algoritmov kde možno spomenúť nástroj EVEGA, ktorého kľúčovou náplňou je vizualizácia grafových algoritmov a ktorý je určený nie len pre študentov ale aj pre vývojárov a programátorov [1].

Ďalšou oblasťou, ktorú by som rád spomenul je využitie vizualizácie pri výučbe teórie automatov. Tu možno ako príklad uviesť nástroj JFLAP [2], ktorý umožňuje tvorbu a simuláciu niekoľkých verzií automatov, vrátane konečných automatov, zásobníkových automatov a jedno a dvoj páskových Turingových strojov [3].

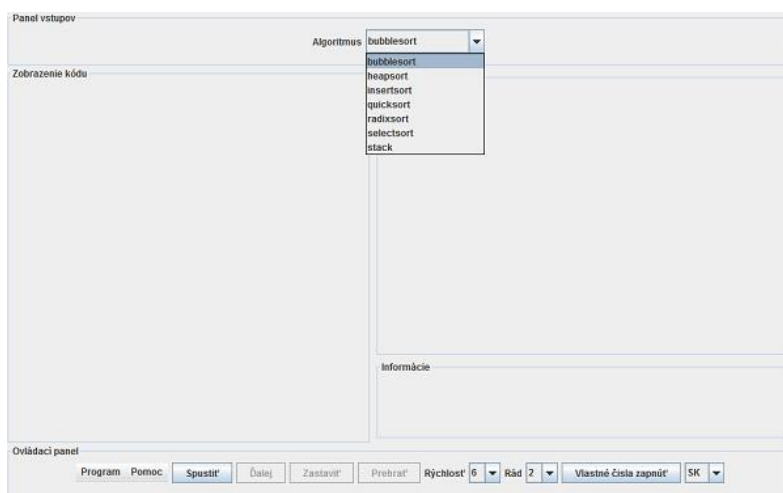
Na záver tejto časti by som ešte rád spomenul oblasť webovej vizualizácie. Konkrétne sa jedna o vizualizáciu údajových štruktúr. Tu možno ako príklad uviesť riešenie JAWAA predstavujúce architektúru pre vytváranie animácií údajových štruktúr bežiacich na webe, ktoré sú písané v jednoduchom skriptovacom jazyku a môžu byť ľahko generované ako výstup programu napísaného v ľubovoľnom programovacom jazyku [4]. Rozsiahlejší prehľad existujúcich riešení možno nájsť v [5].

## III. VIZUALIZAČNÁ PLATFORMA VIZALGO

Jednou z hlavných úloh práce bola analýza a popis súčasného stavu vizualizačnej platformy VizAlgo [6]. Aplikácia je napísaná v programovacom jazyku Java čím je zabezpečená jej prenositeľnosť a podporuje pridávanie nových algoritmov vo forme zásuvných modulov. Samotná aplikácia pozostáva z hlavného programu a z nezávislých modulov obsahujúcich kód daného algoritmu. Na možnosť pridávania zásuvných modulov bol použitý rámec Java Simple

Plugin Framework, ktorý skrýva implementačné detaily komponentov a využíva len ich rozhrania. Jeho základom je využívanie Java notácii a na zavolanie komponentov postačia len dva riadky kódu. Hlavný program rieši samotné ovládanie a zobrazenie daného algoritmu. Pozostáva z niekoľkých tried a rozhraní, ktoré zabezpečujú vzhľad panelov pre vizualizáciu, vykonávaciu logiku, nastavenie daného algoritmu a obsahujú podporné metódy pre správnu funkčnosť aplikácie. Súčasťou hlavného programu sú zásuvné moduly dodané spolu s aplikáciou nachádzajúce sa v adresári *Plugins*. Aplikácia vo svojom základe obsahuje podporu pre vizualizáciu činnosti triediacich algoritmov a operácii na abstraktnom dátovom type zásobník.

Hlavné okno aplikácie (Obr. 1) je tvorené nasledujúcimi oblasťami: panel vstupov, zobrazenie pseudokódu algoritmu, vizualizácia algoritmu, informácie o danom algoritme a ovládací panel. Po načítaní zvoleného algoritmu z panelu vstupov a kliknutí na tlačidlo spustiť sa načíta pseudokód daného algoritmu a čísla, prípadne grafická primitíva na ktorých bude prebiehať samotná vizualizácia. Po načítaní má používateľ možnosť vizualizáciu krokovať, zastaviť alebo prehrať požadovanou rýchlosťou pomocou ovládacieho panelu. Ovládací panel obsahuje ešte možnosť pre voľbu jazyka, nastavenie vlastných čísel a ich rádu ešte pred spustením vizualizácie a tlačidlá pre ukončenie aplikácie a zobrazenie stručných informácií o programe a jeho autorovi [7].



Obr. 1. Používateľské rozhranie aplikácie VizAlgo [7].

#### IV. ROZŠÍRENIE PLATFORMY VIZALGO

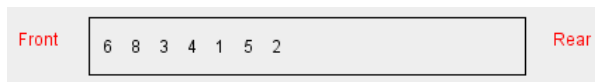
Hlavnou úlohou tejto práce bolo rozšírenie vizualizačnej platformy VizAlgo a to nie len pridaním nových triediacich algoritmov vo forme zásuvných modulov.

Na začiatku samotného rozšírenia aplikácie bol implementovaný nový zásuvný modul pre vizualizáciu algoritmu triedenia so znižujúcim sa prírastkom shellsort a modul pre vizualizáciu operácii na údajovej štruktúre front. V oboch prípadoch bolo potrebné vytvoriť rozhrania obsahujúce dve abstraktné metódy *pseudocode()* a *solution()*, ktoré boli následne implementované v triedach realizujúcich daný modul. Metóda *pseudocode()* zabezpečuje zobrazenie pseudokódu samotného algoritmu a metóda *solution()* obsahuje zdrojový kód samotného algoritmu, ktorý sa bude vizualizovať. Tento kód je možné doplniť pomocnými metódami triedy *ModulViz* ako napríklad vypísanie zotriedených prvkov alebo ich podfarbenie, zvýraznenie riadku zdrojového kódu (Obr. 2) a podobne [8].

```
for (int inc = n / 2;
    inc > 0; inc = (inc == 2 ? 1
    : (int) Math.round(inc / 2.2))) {
    for (int i = inc; i < n; i++) {
        int temp = x[i];
        for (int j = i; j >= inc
            && x[j - inc] > temp; j -= inc) {
            x[j] = x[j - inc];
            x[j - inc] = temp;
        }
    }
}
```

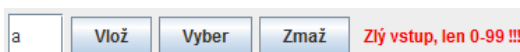
Obr. 2. Zvýraznenie riadku zdrojového kódu algoritmu shellsort.

Čo sa týka modulu pre vizualizáciu operácii na údajovej štruktúre front tak na jeho implementáciu bol použitý obojstranný front *Deque* umožňujúci odstraňovať prvky zo začiatku alebo konca údajovej štruktúry. Samotný front je graficky reprezentovaný (Obr. 3) pomocou obdĺžnika a na vykreslenie jeho aktuálneho obsahu bola vytvorená metóda *printQueue()*.



Obr. 3. Grafická reprezentácia ADT front.

Modul umožňuje pridať nový prvok na koniec frontu, odobrať prvok zo začiatku frontu alebo zmazať celý front. Pridanie nového prvku sa uskutoční až po overení zadaného vstupu, na čo bola vytvorená pomocná metóda *inputTest()* umožňujúca používateľovi nastaviť maximálny počet cifier vstupného čísla. V prípade nesprávneho vstupu sa používateľovi zobrazí text informujúci o jeho nesprávnosti a intervale možných vstupných hodnôt.



Obr. 4. Zadané nesprávne vstupu [7].

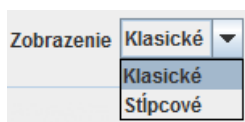
Po dokončení modulov sa vytvorený *.jar* súbor uloží do adresára *Plugins* a po spustení aplikácie sa zobrazí v ponuke dostupných algoritmov.

Vizualizačná platforma VizAlgo vo svojom základe umožňuje vizualizovať činnosť triediacich algoritmov len jedným spôsobom zobrazenia vizualizácie. Cieľom teda bolo umožniť používateľovi zvoliť si spôsob zobrazenia vizualizácie a to medzi pôvodným zobrazením (Obr. 5) pomocou podfarbovania čísel a takzvaným stĺpcovým zobrazením vizualizácie kde je každé číslo reprezentované prostredníctvom stĺpca kde veľkosť daného stĺpca závisí od veľkosti daného čísla.



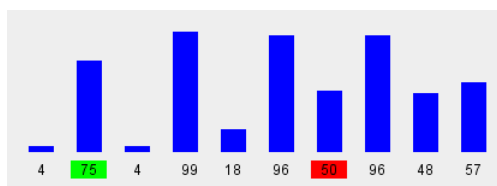
Obr. 5. Pôvodné zobrazenie vizualizácie činnosti triediaceho algoritmu.

Na realizáciu tohto rozšírenia bolo potrebné pridať nového vizuálneho prvku (Obr. 6) do triedy *VizApplet*, ktorá zabezpečuje a vytvára vzhľad panelov pre vizualizáciu a prislúchajúcich komponentov, umožňujúceho používateľovi možnosť voľby zobrazenia vizualizácie.



Obr. 6. Voľba zobrazenia vizualizácie [7].

Ďalej bolo nutné upraviť metódu *draw()* triedy *ModulViz* kde sa na základe zvoleného zobrazenia vykreslia čísla alebo stĺpce. Keďže stĺpcové zobrazenie (Obr. 7) je určené len pre moduly triediacich algoritmov, má používateľ možnosť povoliť alebo zakázať voľbu zobrazenia vizualizácie pre daný zásuvný modul nastavením hodnoty statickej premennej *enableSelectView* triedy *VizApplet* (1 - áno / 0 - nie).



Obr. 7. Stĺpcové zobrazenie vizualizácie činnosti triediaceho algoritmu.

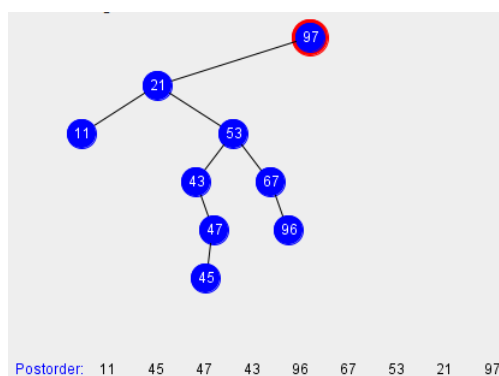
Vizualizačná platforma VizAlgo vo svojom základe a po pridaní modulov spomínaných v tejto časti podporuje len vizualizáciu triediacich algoritmov a abstraktných dátových typov FIFO a LIFO. Cieľom nasledujúceho rozšírenia bolo pridať podporu pre stromy. Konkrétne sa jedná o modul vizualizujúci prechody binárnym a modul na konštrukciu binárneho vyhľadávacieho stromu.

Pri tvorbe modulu pre prechod stromom bolo nutné vytvoriť triedu reprezentujúcu uzol stromu a obsahujúcu metódy slúžiace na vkladanie a mazanie prvkov. Každý uzol stromu si okrem svojej hodnoty uchováva aj svoju x-ovú a y-ovú súradnicu v strome, ktoré nadobudne po vykreslení stromu a ktoré slúžia na jeho označenie počas vizualizácie prechodu stromom. V ďalšom kroku bolo samozrejme nutné upraviť metódu *draw()* o pridanie podpory pre vykreslenie stromu a vytvoriť pomocnú metódu *loadTree()* ktorá zabezpečí rekurzívne vykreslenie stromu prechodom preorder. Voľba vizualizácie konkrétneho prechodu (Obr. 8) stromom je daná indexom položky v zozname prechodov, ktorý sa objaví po načítaní tohto modulu zo zoznamu algoritmov.



Obr. 8. Voľba prechodu binárnym stromom [7].

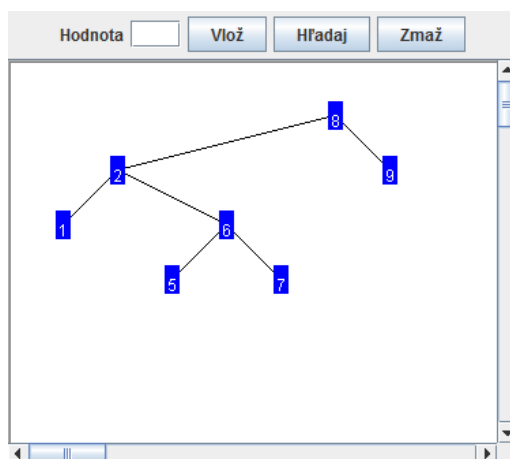
Pri vizualizácii prechodu stromom (Obr. 9) sú jednotlivé uzly označované pomocou pomocnej metódy *treeColorMarker()* triedy *ModulViz* a pod samotným stromom sa súčasne vypisujú hodnoty uzlov v poradí prechodu stromom.



Obr. 9. Vizualizácia postorder prechodu stromom.

Modul na konštrukciu binárneho vyhľadávacieho stromu (Obr. 10) umožňuje používateľovi vkladať, mazať a vyhľadávať prvky v strome. Každý uzol stromu ukladá odkazy na jeho pravého potomka, ľavého potomka a rodiča. Samotné vykresľovanie stromu je realizované v triede *BinaryTreePanel*. Po spustení modulu sa používateľovi otvorí nové okno obsahujúce tlačidlá na vykonávanie spomínaných operácií a vstupné pole, ktoré je ošetrené pomocou metódy *inputTest()*. Samotné okno so stromom sa zavrie až po zastavení činnosti modulu v hlavnom okne aplikácie a teda jeho zatvorenie klasickým spôsobom je používateľovi znemožnené.

Detailnejší popis jednotlivých rozšírení a tried možno nájsť v bakalárskej práci a systémovej príručke na túto tému [7].



Obr. 10. Konštrukcia binárneho vyhľadávacieho stromu.



## V. ZÁVER

Pridaním zásuvných modulov shellsort a queue bolo používateľovi umožnené lepšie pochopenie činnosti algoritmu so znižujúcim sa prírastkom a operácii na abstraktnom dátovom type front. S týmto rozšírením bola zároveň pridaná aj možnosť ošetrovania vstupných údajov od používateľa, čo umožní tvorcom zásuvných modulov ošetriť požadované vstupné pole a určiť interval akceptovateľných vstupov od používateľa. Pridanie novej formy zobrazenia dáva používateľovi možnosť voľby grafickej reprezentácie vizualizácie činnosti triediacich algoritmov a umožňuje mu sledovať vizualizáciu daného algoritmu prostredníctvom stĺpcov, ktorých výška závisí od veľkosti daného čísla. Implementovaním podpory pre stromy bolo umožnené používateľovi pochopenie prechodov binárnym stromom a operácii na binárnom vyhľadávacom strome. Rozšírením aplikácie o tieto zásuvné moduly bolo používateľom umožnené pochopenie širšej oblasti algoritmov a údajových štruktúr a tým teda boli splnené ciele stanovené na začiatku jej rozšírenia. Keďže nezávislosť aplikácie na platforme je zabezpečená voľbou jazyka Java, je spustiteľná na operačných systémoch Linux a Windows za podmienky, že je v systéme nainštalovaná podpora pre beh Java aplikácii. Na iných operačných systémoch aplikácia nebola testovaná. V budúcnosti by bolo ešte možné aplikáciu rozšíriť o podporu pre vizualizáciu grafov, možnosť krokovať vizualizáciu vzad, možnosť dynamických zmien kódu algoritmu pomocou vstupných polí pre používateľa alebo o pridanie nových foriem zobrazenia vizualizácie. Taktiež by bolo vhodné vylepšenie vzhľadu a sprehľadnenie zdrojového kódu aplikácie.

## POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Centrum informačných a komunikačných technológií pre znalostné systémy (ITMS 26220120020) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

## LITERATÚRA

- [1] KHURI, S. – HOLZAPFEL, K.: EVEGA: An Educational Visualization Environment for Graph Algorithms. In Proceedings of the 6th Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE 2001. ACM Press, 2001.
- [2] GRAMOND, E. – RODGER, S. H.: Using JFLAP to Interact with Theorems in Automata Theory. SIGCSE Bulletin (ACM Special Interest Group on Computer Science Education), 31, 1999.
- [3] JFLAP Tutorial. [online] [citované 30. 10. 2013]. Dostupné na Internete: <<http://www.jflap.org/tutorial/>>.
- [4] The JAWAA Homepage. [online] [citované 18. 12. 2013]. Dostupné na Internete: <<http://www.cs.duke.edu/csed/jawaa2/jawaaeditor/jawaa/DOCS/overview.html>>.
- [5] ŠIMOŇÁK, S.: Algorithm Visualization Using the VizAlgo Platform, Acta Electrotechnica et Informatica, Vol.13, No.2 (2013).
- [6] SAJKO, A.: Vizualizácia Algoritmov, FEI TUKE, Bakalárska práca, 2012.
- [7] KUKĽA, L.: Vizualizácia činnosti algoritmov, FEI TUKE, Bakalárska práca, 2014.
- [8] ŠIMOŇÁK, S.: Using Algorithm Visualizations in Computer Science Education, Central European Journal of Computer Science (CEJCS), Versita, ISSN:1896-1533 (prijaté na publikovanie).

# Co-existence of DVB-T/2 and LTE800

<sup>1</sup>František DELY, <sup>2</sup>Lukáš SENDREI, <sup>3</sup>Stanislav MARCHEVSKÝ

<sup>1,2,3</sup>Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>frantisek.dely@student.tuke.sk, <sup>2</sup>lukas.sendrei@tuke.sk,  
<sup>3</sup>stanislav.marchevsky@tuke.sk

**Abstract** — Frequency spectrum that was previously reserved for TV broadcasting was freed up for the use in new mobile radio standard. Now, numerous of co-existence scenarios are possible. The LTE (Long Term Evolution) has opportunity to allocate the spectrum between 790-862 MHz which is part of the frequency spectrum used by DVB-T (Digital Video Broadcasting – Terrestrial). Those two different technologies are now operating on adjacent frequency bands. Based on this fact, it is important to evaluate the amount of interference between DVB-T and LTE and how these interferences impact DVB-T services. The fundamentals of DVB-T systems are described in this paper. Main differences between DVB-T and LTE and also possible solutions to the issue of co-existence are mentioned.

**Keywords** — DVB-T, DVB-T2, LTE co-existence, UHF, Adjacent frequency bands

## I. INTRODUCTION

Most of the European Countries have already finished their transition to Digital Television (DTV) from analogue broadcasting and also it was completed in Slovakia on 31 December 2012. In Slovakia, all terrestrial TV services are now being delivered digitally using Digital Video Broadcasting – Terrestrial technology. These digital services, like the superseded analogue services, use Ultra high frequency (UHF) spectrum for their transmission. The ITU World Radio Conference in 2007(WRC-07) decided to allocate the upper part of UHF spectrum (470-862 MHz), from 790-862 MHz, for mobile services on a co-primary basis with broadcasting. The mobile telecommunication system has opportunity to allocate the spectrum between 790-862 MHz. It means that it is necessary to re-organize spectrum allocation by national authorities over the whole world. There comes the main issue of this paper, because in many countries the mobile telecommunication system 3GPP Long Term Evolution (LTE) transmissions are spectrally adjacent to DVB-T. The interference occurred from mobile system operating on channels 61 to 69 (790 and 862 MHz band) into DVB-T operating on channels up to channel 60. This interference can occur up to the 9 adjacent channel (N+9). There is a 1MHz guard band between 790 and 791MHz, which divides the top of the digital terrestrial TV (DTT) spectrum and the first long-term evolution channel[1][2].

The second issue is that some current LTE channels are still used by existing DVB-T receivers which were strictly designed to use them. DVB-T devices and the corresponding antenna are designed to receive signals in the 800 MHz band. A too strong LTE-800 signal can interfere and affect the DVB-T signal or even superimpose it. On the other way, LTE receivers can receive RF signals in their channels only. There are a lot of possible unwanted effects, like a frozen picture, really bad sound and also complete loss of the signal, in the worst case. Compare to analogue broadcasting, interferences when using digital broadcasting appear suddenly and can appear through a insignificantly degradation of the reception situation. Interferences can arise from irradiations of LTE-base stations and LTE-end devices (e.g. a USB surf stick, modem or mobile phone). Mostly the interference of the end devices is more problematic because it's nearer to the DVB-T recipient. In addition a LTE-end device does not constantly send data [3][4].

In this paper, we will give an overview of the issue of co-existence between DVB2 and LTE800 systems. Based on the current technical specifications, we propose some possible solutions to overcome this particular problem. One solution can be based on either base station (BS) signal filtering or DVB-T receiver filtering. Other option is to reduce the transmitted base BS signal power level. The last solution, mentioned in this paper, is the usage of different polarizations of the BS and DVB-T antennas to achieve additional attenuation.

This paper is organized as follows. Section II describes the fundamentals of DVBT-T/2 and also appears on difference between them. Section III describes fundamentals of LTE. In section IV is described the problem of interference between DVB-T and LTE. The possible solutions are mentioned in section V. The last section gives the conclusions.

## II. DIFFERENCE BETWEEN DVB-T AND DVB-T2

### A. Digital Video Broadcasting –Terrestrial

DVB-T standard was set in 1997 and one year later was put into use for the transmission of digital terrestrial television. DVB-T means using radio frequency spectrum much more efficiently and transmitting digital signal. There are so many advantages over the old, analog broadcasting for that system. It provides an advanced method of transmission compared to the previous analog transmission. This system is able to transmit different kinds of data, including compressed digital information, digital audio, digital video and other data in an MPEG transport stream, using OFDM (Orthogonal frequency-division multiplexing) modulation.

The transmission technique OFDM works by splitting the digital data stream into a large number of slower digital streams, each of which digitally modulate a set of closely spaced adjacent carrier frequencies. There are two choices for the number of carriers known as 2K-mode or 8K-mode for DVB-T and offers 16QAM, 64QAM, QPSK modulation schemes. DVB-T has been adopted or proposed for digital television broadcasting by many countries, using mainly VHF 7 MHz and UHF 8 MHz channels. The first generation considered primarily the standard definition (SD) digital television broadcasting. Data-casting services and high-definition (HD) television content was available via satellite and cable but, until now, terrestrial users have not received such content because its transmission requires wider data channel than for SD. It is possible to transmit data and HD using even current standards – DVB-T for transmitting and MPEG-2, MPEG-4 and H.264 for video coding. Experiences in several countries indicate that terrestrial HDTV services are possible even without using the latest DVB-T2 standard, but not in enough effective way [5][6].

### B. DVB-T2

Research in transmission technology has continued and new option for error-protecting broadcast steams and also new types of modulating have been developed since the DVB-T publication. The newest 2<sup>nd</sup> generation standard for the terrestrial broadcasting is DVB-T2 which offering significant benefits compared to DVB-T. This upgrade enables new high-quality services, such as the HDTV, through terrestrial networks. It is motivated by higher spectral efficiency, which means that the same number of programmers broadcast with higher audio and also video quality or larger number of programmers can be broadcast at all. There is improved source of coding MPEG-4 and its gain in broadcast transmission is remarkable. Also much more TV- programs in HD quality become possible with this technology. DVB-T2 boasts up to 256QAM modulation, plus higher number of carriers to deliver much more robust signal.

DVB-T2 uses LDPC (low density parity check) and also BCH (Bose-Chaudhuri-Hocquengham) coding for error protection. There is also a new technique, known as Rotated Constellations, has been introduced to provide additional robustness in certain conditions.

DVB T2 standard takes also care of transmitter equipments. In 32k particularly, high power peaks are generated and thus minimize the amplifier efficiency (or even undamaged it). A special feature called PAPR (Peak Average Power Ratio) reduction has been included in the standard specifications to limit these peaks power without losing information [7][8][9].

### C. Main Differences between DVB-T and DVB-T2

**Tab. 1 DVB-T vs DVB-T2**

Main Differences:	DVB-T	DVB-T2
FEC:(Forward Error Correction)	Conv.Coding +Red-Solomon	LDPC + BCH
Code rate:	1/2, 2/3, 3/4, 5/6, 7/8	1/2, 3/5, 2/3, 3/4, 4/5, 5/6
Modes:	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM, 256QAM
Guard Interval:	1/4, 1/8, 1/16, 1/32	1/4, 19/256, 1/8, 19/128, 1/16, 1/32, 1/128
Bandwidth:	6, 7, 8 Mhz	1.7, 5, 6, 7, 8, 10 MHz
Capacity:	31,66 Mb/s	50,34 Mb/s

### III. LONG TERM EVOLUTION – LTE

LTE is radio platform technology that allows operators to achieve even higher peak throughputs than HSPA + (High Speed Packet Access) in higher spectrum bandwidth. It is actually the latest standard in the mobile network technology realized by 3<sup>rd</sup> Generation Partnership Project (3GPP). Work on this technology began in 2004 and was completed by 3GPP Release 8 in March 2009. LTE is a part of the GSM (Global System for Mobile Communications) evolutionary path for mobile broadband, following EDGE, HSPA, HSPA+ Even though 3GPP created standards for GSM/UMTS family, the LTE standards are completely new, with exceptions where it made sense. The over objective for LTE is to provide an extremely high performance radio-access technology that offers full vehicular speed mobility and that can very easy coexist with HSPA and earlier networks and also with others systems.

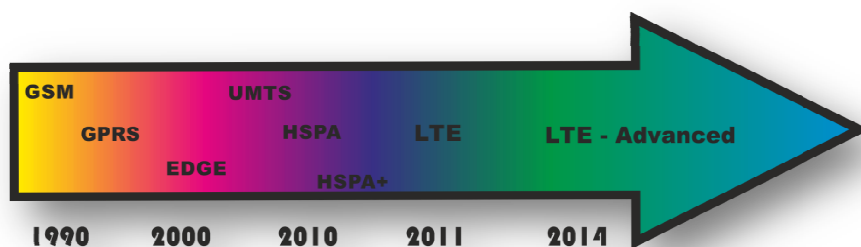


Fig. 1 3GPP Family Technology Evolution

This system called fourth generation (4G) provides users with much faster data speed than 3G is able to and it is easy deployable network technology offering except high speed, low latencies over long distances. It uses two different types of air interfaces (radio links), one for downlink from tower to device and one for uplink from device to tower. LTE also utilizes the optimal way to do wireless connections both ways, which makes a better optimized network. For the downloading it uses Orthogonal Frequency Division Multiple Access (OFDMA), which is well suited to achieve high peak data rates in high spectrum bandwidth. It uses an approach for the uplink called DFTS-OFDMA (Discrete Fourier transform spread orthogonal frequency division multiple access) scheme of generating a SC-FDMA (Single Carrier Frequency Division multiple access) signal. SC-FDMA is better for the uplink than OFDMA, because it has a better peak-to-average power ratio for uplink. The major difference between them is that SC-FDMA uses a discrete Fourier transform function on the data to convert it into a form that can be used to transmit. That Fourier functions are often used to convert digital data into analog waveforms for decoding audio and video and also for outputting the proper audio frequencies [8][9][10].

### IV. CO-EXISTENCE –THE PROBLEM OF LONG-TERM EVOLUTION

#### A. 4G frequency bands

- The 2.6 GHz band is one of the two frequencies that were auctioned off by Ofcom (communications regulator) for LTE. That band has a greater data capacity than the other one, so it can deal with loads of people connection at once, but the problem is, that it doesn't fare so well over long distances, making it ideal for cities and other compact, densely populated areas but not so good for rural locations.
- The second one is the 800 MHz band. That band was used to provide analogue terrestrial TV, but has been freed up since the big Digital switchover. It is true, that it doesn't provide the same capacity as the 2.6 GHz band, but this one an easily travel over long distances and will be used to provide broadband speeds to rural areas where telephone exchanges can't reach [11].

#### B. The frequency bands problem

TU-R WRC 2007 dedicated to allocate 72 MHz, in the frequency range 790-862 MHz in region 1 for mobile services on a co-primary basis, which means that used TV channels 61-69 should no longer be allocated to broadcasting. Broadcasting services are now limited to the frequency band 470 – 790 MHz, ie. to broadcasting channels 21-60. Deploying mobile services in the frequency range 790-862 MHz create a potential risk of adjacent channel interference into existing digital TV services using channels below channel 61[4].

As it was mentioned, LTE uses frequencies belonging to the sub-band from 791 to 962 MHz. For the downstream: 791 MHz to 821 MHz and 832 MHz to 862 MHz for the upstream. There is just small separation (1MHz) between the band reserved for TV and LTE.

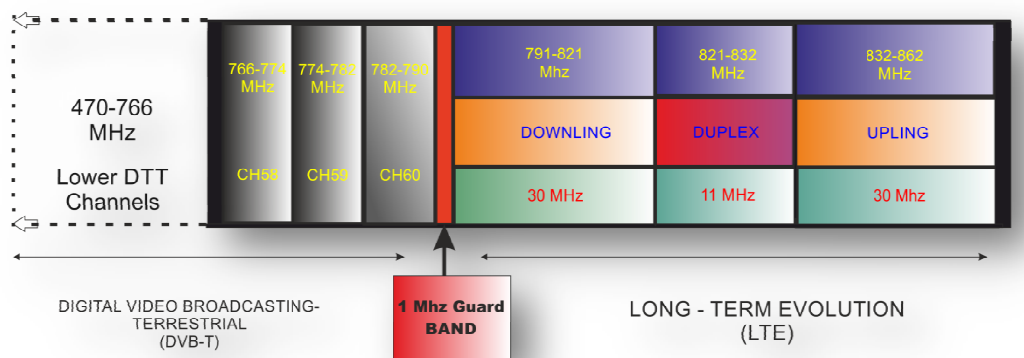


Fig. 2 Harmonized 800 MHz digital dividend spectrum organization

The close proximity of cellular and digital TV assignments will result in interference to digital TV services in some areas, particularly near the edge of coverage. This will require co-ordination between the broadcasters and the mobile operators to manage the problem. The mobile service uplink as well as the downlink may interfere with DVB-T. The network structure for DVB-T consists of a relatively small number of sites often using high output power in combination with high antenna heights. The mobile network has a denser network structure, with a higher number of base stations and lower power levels.

### C. Receivers of unwanted signals

Both receivers of (DVB-T or LTE) systems receive unwanted signals of other radio systems or unwanted parts such as noise too. DVB-T receivers that are located in areas close to mobile network base stations are likely to receive relatively high field strength from mobile network base stations, compared to the field strength from DVB-T transmitter as it shown on figure 4. There are various coexistence scenarios where the signal of one system may act as an unwanted signal to the respective other system. The severity of interference depends on degrading of receiver quality of digital TV reception or even loss of throughput in case of mobile radio.

There are two different coexistence scenarios for DVB-T and LTE possible and can be summarized as follows [2][12][13]Chyba! Nenalezen zdroj odkazů.:

- DVB-T influenced by Mobile Radio  
 Here, the DVB-T transmitter may act as interferer on the cellular system LTE. According the type of spectrum simulation, LTE terminal receiver or its base station receiver could be impacted. As the DVB-T and also LTE systems are operated in the different frequency bands, that kind of coexistence will never be a co-channel scenario.
- LTE influenced by DVB-T  
 In this case, an LTE transmitter acts as interferer on DVBT-T receiver (settop box). The LTE transmitter can be either an LTE terminal transmitting an uplink signal, or an LTE base station transmitting a downlink signal.

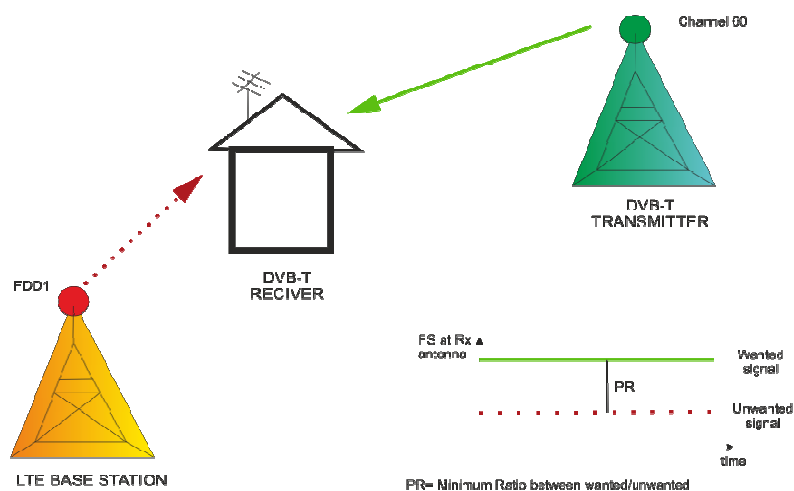


Fig. 3 The unwanted signals are received and interferes with receiver preventing it from decoding the wanted signal correctly

## V. POSSIBLE SOLUTIONS FOR INTERFERENCE MITIGATIONS

Numbers of mitigation measures that can be used to restore DVB-T services or mitigate the risk of receiver overload are. In this section, we describe a range of possible mitigation techniques that can be envisaged, and consider the effect of each technique individually.

### A. Base station filtering

At the base station can be effective to reduce the out of band emissions if it there are dominating the overall receiver performance. Reducing the BS out of band emissions enables benefit of fitting an external filter to a DTT receiver with a good adjacent channel protection radio.

### B. Built in filtering antennas

Indoor portable antennas or roof top antennas are often designed to have a natural roll-off a 790MHz without the use of active circuitry, which helps to reduce the amount of interference the input of booster amplifiers of TV tuners. There is one disadvantage occurred by existing antenna, because it needs to be replaced.

### C. End receiver external filtering

One of the most effective solutions is adding a low pass inline filter between the TV tuner input and TV antenna when the highest TV channel is below channel 59 or 60. Before the first active component in the receiving chain, usually a broadband amplifier should be any filter added. Reducing the filter cut-off frequency below 790 MHz increases the amount of LTE interference rejection at the cost of increasing tilt and insertion loss across the wanted channels below 790 MHz, which operating at the edge of coverage, could result in loss of reception.

### D. BS power level reduction

Another way of mitigate interference from LTE BS, particularly for the worst case interferer of LTE in areas where is the biggest impact and a lot of homes are affected. An adjustment of a BS antenna in order to reduce BS power level into a certain direction or area could be treated as a sub-category of this method.

### E. Polarization of antenna

It uses different polarizations of the BS and DVB-T signals to achieve additional rejection of the LTE BS signal in the DVB-T receiver if the signal is received via a roof top aerial. However, the use of any polarization for mitigation purposes is limited to fixed DVB-T reception, only

Other mitigation techniques can also be applied, to improve the DVB-T signal strength with respect to the LTE signal strength, thereby reducing the impact of interference. In order to minimize the interference from the mobile network it seems necessary to insert a frequency separation (guard band) of 2 MHz above 790 MHz. This will reduce the impact of interference to less than 8% of the population where DVB-T channel 60 is received [4], [1].

## ACKNOWLEDGMENT

This work is the result of the project implementation: Development of the Center of Information and Communication Technologies for Knowledge Systems (ITMS project code: 26220120030) supported by the Research & Development Operational Program funded by the ERDF.

## VI. CONCLUSION

This article was focused on possible problems between two different technologies which are operating on adjacent frequency bands. As we have found out, that the biggest problem is the just 1MHz guard band. A lot of TV channels has problem with signals influenced by LTE system which got a new frequency spectrum for use. Normally, mobile bands are planned with the *uplink* (base station receive/mobile transmit) in the lower band and the *downlink* (base station transmit/mobile receive) in the upper band. However, due to concerns regarding interference from future mobile transmission to DTT below 790MHz, it was decided for the 800MHz band to reverse the duplex direction, so that the *downlink* is in the lower band. There are possible mitigation techniques to restore DVB-T/2 services in the presence of interference. One of the most obvious solutions is filtering, although improvements to DVB-T receiver specifications over time could also provide an equally effective means of mitigating interference.

#### REFERENCES

- [1] DotEcon Ltd: 800MHz auction: Co-existence of LTE systems in 790-862 MHz with Digital Terrestrial Television, August 2011.
- [2] Sakic, K. and Grgic S., "The Influence of the LTE System on DVB-T Reception", September 2010, Zadar, Croatia.
- [3] Toshiba: Interference to TV signals (Long Term Evolution- LTE) Online: <<http://aps2.toshiba-tro.de/kb0/TSB21038P0000R01.htm>>.
- [4] PROGIRA Radio Communication: Interference from future mobile network services in frequency band 790-862 MHz to digital TV in frequencies below 790 MHz, May 2009, Online: <<https://www.pts.se/upload/Rapporter/Radio/2009/Report-DVB-T-vs-mobile-network-interference-study-2008-02-05.pdf>>.
- [5] TECHNOPEdia, Cory Janssen: Digital Video Broadcasting- Terrestrial (DVB-T) Online: <<http://www.techopedia.com/definition/25174/digital-video-broadcasting-terrestrial-dvb-t>>
- [6] I. Funa, " DVB-T ", Online: < <http://igorfun.com/dvb-t/>>.
- [7] Reljin, IS.; Sugaris, AN., "DVB standards development," Telecommunication in Modern Satellite, Cable, and Broadcasting Services, 2009. TELSIKS '09. 9th International Conference on , vol., no., pp.263,272, 7-9 Oct. 2009.
- [8] Enensys technologies: General overview of DVB-T2 standard, Online: <<http://www.enensys.com/technologies/dvb-t2-overview.html>>.
- [9] 4G: LTE: Long Term Evolution, August 2013 Online: <<http://www.4gamericas.org/index.cfm?fuseaction=page&sectionid=249>>.
- [10] MOTOROLA: Long Term Evolution (LTE), 2007, Online: <[http://www.motorolasolutions.com/web/Business/Solutions/Industry%20Solutions/Service%20Providers/Wireless%20Operators/LTE/\\_Document/Static%20Files/6833\\_MotDoc\\_New.pdf](http://www.motorolasolutions.com/web/Business/Solutions/Industry%20Solutions/Service%20Providers/Wireless%20Operators/LTE/_Document/Static%20Files/6833_MotDoc_New.pdf)>.
- [11] ExtremeTech, NEAL GOMPA: What is LTE?, February 2013, Online: <<http://www.extremetech.com/mobile/110711-what-is-lte/7>>.
- [12] J. Rogerson, "4G and LTE: everything you need to know", December 2013, Online: <<http://www.techradar.com/news/phone-and-communications/mobile-phones/4g-and-lte-everything-you-need-to-know-926835>>.
- [13] Rohde & Schwarz GmbH & Co.KG, BERNHARD SCHULZ, Coexistence Digital TV and LTE, Augustu 2010.
- [14] Baruffa, G.; Femminella, M.; Mariani, F.; Reali, G., "Protection Ratio and Antenna Separation for DVB—T/LTE Coexistence Issues," Communications Letters, IEEE , vol.17, no.8, pp.1588,1591, August 2013.

# Kognitívne rádio

<sup>1</sup>Peter ČARNAKOVIČ, <sup>2</sup>Lukáš SENDREI, <sup>3</sup>Stanislav MARCHEVSKÝ

<sup>1,2,3</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>peter.carnakovic@student.tuke.sk, <sup>2</sup>lukas.sendrei@tuke.sk,  
<sup>3</sup>stanislav.marchevsky@tuke.sk

**Abstrakt** — Kognitívne rádio (Cognitive Radio) je forma bezdrôtovej komunikácie, v ktorej vysielateľ môže inteligentne rozpoznať, ktoré komunikačné kanály a frekvenčné pásma sa používajú a ktoré nie. Vďaka tomu kognitívne rádio umožňuje použitie licenčných frekvenčných pásiem aj sekundárnym užívateľom, ktorí nemajú licenciu. Avšak títo sekundárni (kognitívni) užívatelia potrebujú neustále sledovať frekvenčné spektrum, aby sa tak zabránilo možnému rušeniu primárnych užívateľov. Tento článok podáva krátky úvod do problematiky kognitívnych rádii.

**Kľúčové slová** — kognitívne rádio, softwarové rádio, SDR, DSA, CR, snímanie spektra, biele miesta

## I. ÚVOD

Bezdrôtová technológia sa rýchlo rozrastá do všetkých aspektov výpočtovej techniky a komunikácie. V dnešnej dobe existuje viac ako 3 mld. bezdrôtových zariadení (sú to predovšetkým mobilné telefóny a prenosné počítače), ale do roku 2025 sa očakáva, že toto číslo bude predstavovať až 100 mld. bezdrôtových zariadení [1]. Preto bude v budúcnosti rádiová technológia tvoriť jadro celého výpočtového sveta, v ktorom budú miliardy komunikátorov, mobilných zariadení a senzorov pripojených na internet a bude tvoriť základ pre mnoho vzrušujúcich nových tried aplikácií [2], [3]. Tento očakávaný exponenciálny rast bezdrôtových zariadení a aplikácií je závislý na našej schopnosti navrhnuť a vytvoriť rádiové technológie, ktoré budú aj naďalej schopné dobre spolupracovať pri zvyšujúcej sa hustote nasadenia – najmä sa musia zmeniť rádiové systémy, a to veľmi rýchlo, aby sa prispôbili náhlemu nárastu rozsahu hustoty z dnešných 10-100 zariadení/km<sup>2</sup> na 1000-10,000 zariadení/km<sup>2</sup> v roku 2025. Očakávaný rozmach bezdrôtových aplikácií vytvára stále rastúci dopyt po rádiovom spektre. Kvôli tomu vyžaduje oblasť rádiovkej technológie veľkú inováciu vzhľadom na skutočnosť, že spektrum je obmedzené a vzácny zdroj. Frekvenčné pásma boli v minulosti na rôzne účely pridelené staticky [4]. Niektoré časti frekvenčného spektra boli pridelené aplikáciám, ktoré sú dôležité pre spoločnosť ako napr. bezpečnostné a záchranné služby, prvá pomoc a letectvo. Ďalšie frekvenčné pásma boli pridelené na komerčné použitie, napr. mobilné telefóny, rozhlasové a televízne vysielanie. Napokon, obmedzené množstvo spektra je bezlicenčné a môže sa používať (v závislosti na určitých obmedzeniach) ktokoľvek [5]. Je dosť zaujímavé, že práve niektoré bezlicenčné pásma sú najviac používané, pretože pre používateľa je oveľa jednoduchší a lacnejší prístup práve k týmto pásmam. Statické pridelenie frekvenčných pásiem má veľa výhod. Prvou výhodou je jednoduchosť: pásma boli pridelené raz, teda nie je tam žiadna dvojznačnosť o tom, kto môže používať spektrum a presadzovanie predpisov je pomerne jednoduché. Po druhé, rádia boli v minulosti pevne funkčné zariadenia, ktoré boli navrhnuté pre určité frekvenčné pásmo. V dôsledku toho, použitie špecifického frekvenčného pásma vyžaduje značné investície do infraštruktúry, ktoré môžu byť použité len v tomto pásme. Takáto investícia má zmysel iba v prípade, že potrebujeme záruku nepretržitého prístupu k tomuto pásmu. Použitie takýchto vyhradených pásiem nielen zjednodušuje nasadenie niektorých služieb, ale často vedie aj k lepšej kvalite služieb. Dôvodom je, že jeden správca môže riadiť celé nasadenie a fungovanie systému, bez toho, aby mal obavy o ďalších ľubovoľných konkurenčných užívateľoch. Zatiaľ čo statické pridelenie frekvenčných pásiem má mnoho výhod, môže viesť aj k veľmi neefektívnemu využívaniu spektra. Veľa štúdií ukázalo, že mnohé pridelené frekvenčné pásma sú značne nevyužívané [6]-[8]. Jedným z dôvodov je to, že využívanie spektra je často obmedzené (napr. v okolí letísk, atď.). Kognitívne rádio predstavuje inovatívnu technológiu, ktorá umožní budúci bezdrôtový svet.



Predkladaný článok podáva teoretický úvod do problematiky kognitívneho rádia. Z pohľadu prístupu k frekvenčnému spektru, čo je dnes možné jedine staticky na vopred pridelených kanáloch, je myšlienka dynamického prístupu k frekvenčným zdrojom revolučnou myšlienkou, kde technológia kognitívneho rádia umožní v dohľadnej budúcnosti tieto koncepty aplikovať v reálnej prevádzke. Preto pokladáme za nevyhnutné venovať sa tejto problematike a aj týmto príspevkom predložiť prípadným záujemcom o túto tému základné poznatky a najnovší stav problematiky v tejto veľmi rapídne sa vyvíjajúcej a zaujímavej oblasti bádania komunikačných technológií.

## II. SOFTWARE RÁDIO

Kľúčovou technológiou umožňujúcou zavedenie kognitívnych bezdrôtových sietí je softvérové rádio (Software Defined Radio) [9], [10], ktoré začalo vznikáť v roku 1990 a teraz sa postupne nasadzuje aj v komerčnej sfére. Technológia SDR prináša rádioelektroniku do digitálneho veku [11], čím dostávame nové možnosti pri návrhu bezdrôtových systémov. To znamená, že programovateľné rádia s možnosťou snímania spektra môžu dynamicky zdieľať spektrum s prostredníctvom rôznych systémov a služieb. Tento prevádzkový režim sa všeobecne nazýva dynamický prístup k spektru DSA (Dynamic Spectrum Access) a je zástupcom prvej vlny výskumu, vývoja a komerčnej činnosti kognitívneho rádia. Bezdrôtové systémy s DSA boli podporované predpismi FCC (Federal Communications Commission), ktoré napr. umožňujú zdieľaný prístup spektra v tzv. bielych miestach TV (z angl. TV White Spaces) [12]. Výzvou v zavádzaní princípov DSA je výrazne zlepšiť účinnosť využívania spektra, bez strát výhod, ktoré sú spojené so statickým pridelovaním spektra. Prvou výzvou je vytvoriť bezdrôtové zariadenia a siete, ktoré môžu príležitostne pôsobiť v rôznych frekvenčných pásmach. Ďalšie výzvy sú v oblasti predpisov pridelenia frekvenčného spektra. Je potrebné vytvoriť normy pre dynamický prístup k spektru, ktoré vedú k efektívnemu využívaniu spektra, sú praktické, chránia práva primárnych užívateľov a zabezpečujú kvalitu služieb. Predpisy musia chrániť záujmy primárnych užívateľov, ktorí vykonali veľké investície do infraštruktúry. Okrem toho, musí byť ekonomicky atraktívna pre výrobcov a poskytovateľov služieb pri vývoji a nasadení zariadení pre oportunistický prístup k spektru sekundárnymi užívateľmi. V nasledujúcich rokoch programovateľná SDR technológia umožní vytvoriť adaptívne bezdrôtové siete, v ktorých môžeme dynamicky vyberať protokoly a prispôbovať rádiové signály pre aktuálne prevádzkové požiadavky. Tieto adaptívne siete dosiahnu ešte vyššiu účinnosť využitia spektra a výkon pomocou distribuovaných kognitívnych algoritmov.

### A. Biele miesta v frekvenčných pásmach

Biele miesta boli zavedené väčšinou z technických dôvodov, aby sa zabránilo rušeniu medzi použitými rozhlasovými pásmami a kanálmi – tzv. ochranné pásmo. Okrem bieleho miesta prideleného z technických dôvodov, sú aj nevyužité oblasti rádiového spektra, ktoré nikdy neboli použité, alebo sa stávajú bezlicenčnými v dôsledku technických zmien. Najmä po prechode na digitálne televízne vysielanie sa uvoľnil veľký priestor medzi 50 MHz a 700 MHz. V roku 2008 FCC vydala nariadenie pre biele miesta v televíznom vysielaní [13]. TV na vysielanie používa niektoré kanály v spektre striedavo, takže vynechané kanály môžu byť bezpečne použité znovu. Tieto nevyužité biele miesta v TV vysielaní sú atraktívnym cieľom pre DSA, pretože pracujú na jednoducho prístupných frekvenciách, a ich dostupnosť je pomerne statická. Aj keď na prvý pohľad sa to zdá byť jednoduché v skutočnosti je to oveľa zložitejšie, pretože tieto prístroje, ktoré by využívali tieto biele miesta musia byť schopné určiť, či je kanál obsadený alebo nie, a podľa toho obmedziť energiu na prenos, aby sa tým zabránilo rušeniu televíznych prijímačov [14].

## III. KOGNITÍVNE RÁDIO

Kognitívne rádiá sú plne programovateľné bezdrôtové zariadenia, ktoré dokážu snímať svoje prostredie a dynamicky prispôbovať prenosový signál, metódy prístupu na kanál, využitie spektra a sieťové protokoly potrebné pre dobrý výkon siete a aplikácií. Budovanie a nasadenie siete kognitívneho rádia je zložitá úloha. Výskumná komunita, ktorá sa zaoberá kognitívnymi rádiovými sieťami musí pochopiť širokú škálu problémov týkajúcich sa technológie inteligentných antén, snímania a merania spektra, spracovania rádiového signálu, hardvérových architektúr vrátane softvérovo definovaného rádia SDR (Software Defined Radio), riadenia prístupu k médiu MAC (Medium Access Control), zisťovania siete a sebaorganizácie, smerovania, adaptívneho riadenia mechanizmov, definovania a monitorovania postupu a mechanizmu učenia. Je to veľmi široká škála technológií, ktoré je potrebné aplikovať a využiť. Preto pochopenie správania sa výsledného systému to predstavuje veľmi náročnú úlohu.

### A. Kognitívne rádiové siete

Siete DSA prinášajú celý rad výziev v oblasti bezdrôtových sietí a komunikácií. Prvým a najvýraznejším problémom je, že potrebujeme zariadenia, ktoré môžu pracovať v rôznych frekvenčných pásmach - väčšina dnes používaných zariadení bolo navrhnutých pre konkrétne frekvenčné pásmo. Po druhé, potrebujeme zariadenia, ktoré sú schopné "rozhodnúť", ktoré frekvenčné pásmo si majú zvoliť a používať, pretože vo väčšine prípadov je nereálne, že užívateľ je schopný zvoliť správne pásmo. Tento proces zvyčajne spočíva v dvoch krokoch: zariadenia potrebujú určiť, aké frekvenčné pásma sú k dispozícii, s ohľadom na príslušné predpisy FCC, a potom je potrebné rozhodnúť, ktoré pásmo je najvhodnejšie. Konečnou úlohou je správne nakonfigurovať bezdrôtovú sieť. Bezdrôtové zariadenia sa musia dohodnúť na tom, ako zrealizovať rôzne funkcie fyzickej, linkovej a sieťovej vrstvy takým spôsobom, ktorý umožňuje čo najlepšie využiť dostupné spektrum a zároveň spĺňať predpisy, ktoré platia vo zvolenom pásme. Kognitívne rádiové siete CRN (Cognitive Radio Network) sú veľmi vhodné pre realizáciu vyššie uvedených funkcií. Všeobecne povedané, CRN sú siete, ktoré môžu snímať ich pracovné prostredie a prispôbiť prevádzku siete tak, aby sa dosiahol čo najlepší výkon. Dnešné bezdrôtové siete (napr. WiFi) už používajú obmedzené formy kognitívnej optimalizácie (napr. prispôsobenie prenosovej rýchlosti) a spektra (napr. výber kanála), ale potrebujeme oveľa výraznejšie prispôsobenia, napríklad v rámci použitia širšieho frekvenčného spektra, radikálnejšej optimalizácie bežiacieho protokolu, zlepšenie účinnosti využitia spektra a kapacity bezdrôtovej siete.

### B. Funkcie kognitívneho rádia

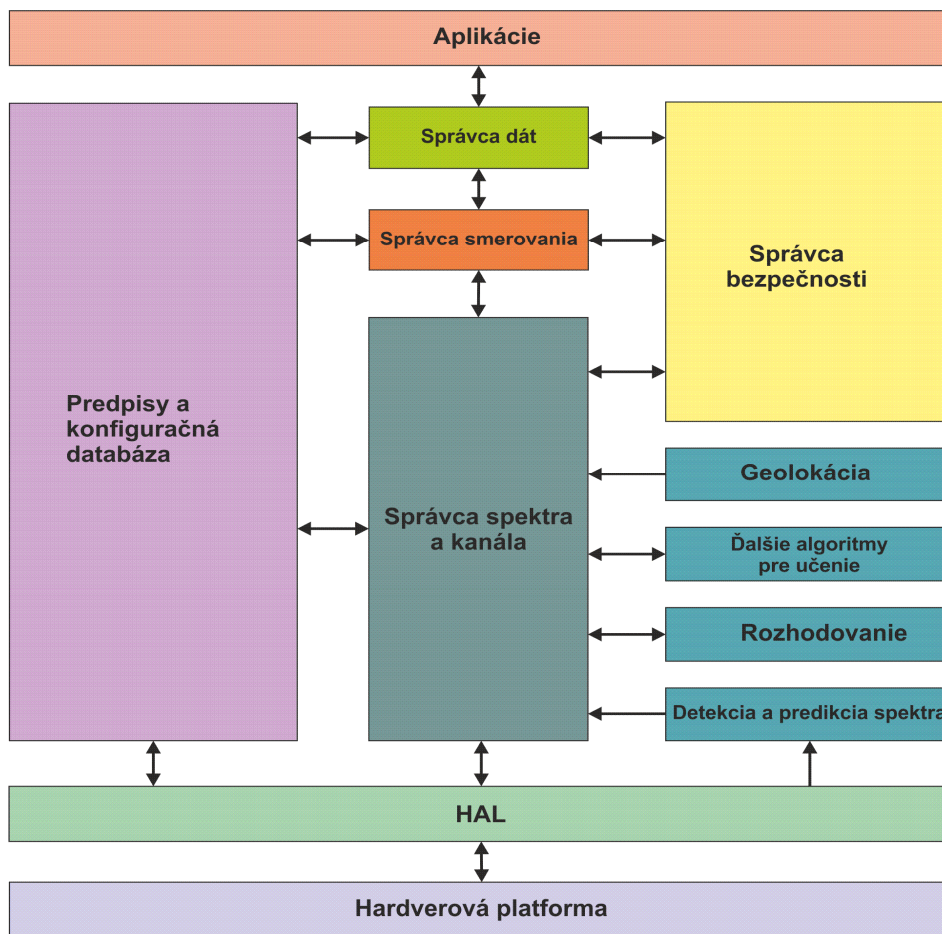
CRN sa môže skladať z radu SDR zariadení, ktoré spolupracujú a dynamicky prispôbujú vlastnosti prenosových vln a signálov k danému prostrediu. Niektoré funkcie CR sú:

- Snímanie aktuálnej frekvenčnej oblasti - Zahŕňa meranie frekvencií, ktoré sa práve používajú. CR sa potom snaží odhadnúť umiestnenie vysielateľov a prijímateľov a zistiť moduláciu signálu na týchto použitých frekvenciách. Výsledky zo snímania prostredia môžu byť použité pre nastavenie parametrov rádia.
- Predpisy a konfiguračné databázy - Predpisy stanovujú fyzické obmedzenia prevádzky rádia v danom spektre a môžu byť uložené v zariadení alebo sprístupnené do celej siete. Predpisy by mohli určiť, ktoré frekvencie môžu byť použité v ktorých miestach. Konfiguračné databázy by mali popisovať prevádzkové charakteristiky fyzického rádia. Tieto databázy by mali byť použité na obmedzenie prevádzky rádia, aby zostali v rámci regulačných alebo fyzických možností.
- Vlastná konfigurácia - Zariadenie môže byť zostavené z niekoľkých modulov ako je anténa, digitálny signálový procesor a riadiaci procesor. Každý modul by mal byť samo-popisný a rádio by tak malo automaticky nakonfigurovať samo seba na prevádzku pomocou dostupných modulov.
- Adaptívne algoritmy - Počas prevádzky kognitívne rádio sníma svoje prostredie, dodržiava predpisy a konfiguračné obmedzenia a komunikuje s ďalšími systémami, aby mohli čo najlepšie využívať rádiové spektrum a spĺňať požiadavky užívateľov.
- Distribúovaná spolupráca - Kognitívne rádiá medzi sebou pravidelne vymieňajú aktuálne informácie o ich prostrediach, užívateľov a rádiových výkonoch. Rádiá používajú svoje miestne informácie a vzájomné informácie pre určenie ich prevádzkového nastavenia.
- Bezpečnosť - Rádiové siete vyžadujú mechanizmy pre autentizáciu, autorizáciu a ochranu informačných tokov účastníkov.

### C. Architektúra

Architektúra uzla kognitívnej rádiovkej siete je zobrazená na Obr. 1. Hardvérová platforma môže byť ľubovoľne taká, ktorá vyhovuje účelom kognitívneho rádia. Hardvérová abstraktná vrstva HAL (Hardware Abstraction Layer) tvorí rozhranie medzi hardvérom a softvérom. Poskytuje dátové rozhranie tiež k dátam prijímateľa a vysielateľa, rovnako ako aj rozhranie na prístup do pamäte. Správca spektra a kanála spravuje všetky prostriedky, ktoré súvisia so spektrom a kanálom, vrátane spojov, frekvenčných a modulačných metód. Správca spektra a kanála tvorí rozhranie s niekoľkými ďalšími funkčnými modulmi. Modul detekcie a predikcie spektra poskytuje informácie, ktoré súvisia s dostupnosťou určitých frekvenčných segmentov. Rozhodovací modul pomocou rozhodovacích algoritmov určuje, ktorý kanál má byť použitý a kedy má byť použitý. Môžeme implementovať aj ďalšie algoritmy učenia, ktoré tvoria nezávislé moduly na učenie od vstupov. Geolokácia určuje zemepisnú šírku a dĺžku uzla. Správca spektra a kanála môže použiť tieto geolokačné informácie na získanie predbežných informácií o aktuálnej polohe z predpisov a z konfiguračnej databázy. Správca smerovania pomocou rôznych

smerovacích algoritmov určuje najlepšiu trasu na prenos dátových paketov. Správca dát organizuje všetky dáta aplikačnej vrstvy, ktoré sú potrebné na prenos. Správca bezpečnosti poskytuje šifrovanie pre správcu dát, správcu smerovania a pre správcu spektra a kanálov. Konfiguračná databáza ukladá predchádzajúce znalosti, predpisy, dáta a skúsenosti [15].



Obr. 1 Architektúra uzla kognitívneho rádia

#### IV. ZÁVER

Rýchly vývoj bezdrôtových technológií spôsobuje čoraz väčší dopyt po rádiovom spektre, ktoré je veľmi limitovaný zdroj. Vznikajúce kognitívne rádiové technológie predstavujú inováciu, ktorá by mohla poskytnúť riešenie škálovania bezdrôtových systémom pre najbližších 25 rokov. Kognitívne rádiové siete predstavujú výrazný posun rádiových a sieťových technológií, ktoré majú veľký potenciál na zvýšenie výkonu systémov a pre efektívne využitie spektra. Stále však existuje veľa technických prekážok, ktoré je potrebné prekonať, aby sme ich mohli realizovať v reálnych scenároch. Najväčšie problémy spôsobujú skrytí primárni užívatelia a rozprestreté spektrum primárnych užívateľov, na základe ktorých sa kognitívne rádio nevie správne rozhodnúť, či daný blok spektra je neobsadený. Nesprávnym rozhodnutím dôjde k vysielaniu signálu na týchto kanáloch, čo následne vedie k rušeniu primárnych užívateľov. Našťastie existuje riešenie týchto, ako aj ďalších problémov, ale je potrebné zdokonaľiť hardvér na snímanie spektra a vytvoriť centrálnu databázu súčasných primárnych užívateľov.

#### POĎAKOVANIE

Táto práca bola podporovaná Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a Slovenskej akadémie vied (VEGA) na základe projektu č. 1/0766/14.

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS//projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

POUŽITÁ LITERATÚRA

- [1] Wireless 2025: The Future of Mobile, Telephony, April 2009. Dostupné na internete: <http://telephonyonline.com/>
- [2] Mitola, J.; Maguire, G.Q., Jr., "Cognitive radio: making software radios more personal," *Personal Communications, IEEE*, vol.6, no.4, pp.13,18, Aug 1999.
- [3] Haykin, Simon, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol.23, no.2, pp.201,220, Feb. 2005.
- [4] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge University Press, 2009.
- [5] National Research Council. *Handbook of Frequency Allocations and Spectrum Protection for Scientific Uses*. Washington, DC: The National Academies Press, 2007.
- [6] Federal Communications Commission, "Spectrum Policy Task Force," Rep. ET Docket no. 02-135, Nov. 2002.
- [7] M. McHenry, "Frequency agile spectrum access technologies," in *FCC Workshop Cogn. Radio*, May 19, 2003.
- [8] Staple, G.; Werbach, K., "The end of spectrum scarcity [spectrum allocation and utilization]," *Spectrum, IEEE*, vol.41, no.3, pp.48,52, March 2004.
- [9] Mitola, J., III, "Software radios: Survey, critical evaluation and future directions," *Aerospace and Electronic Systems Magazine, IEEE*, vol.8, no.4, pp.25,36, April 1993.
- [10] Buracchini, E., "The software radio concept," *Communications Magazine, IEEE*, vol.38, no.9, pp.138,143, Sep 2000.
- [11] Koch, P. and Prasad, R., *The Universal Handset*, *IEEE Spectrum*, April 2009.
- [12] Harrison, K.; Mishra, S.M.; Sahai, A., "How Much White-Space Capacity Is There?," *New Frontiers in Dynamic Spectrum*, 2010 *IEEE Symposium on*, vol., no., pp.1,10, 6-9 April 2010.
- [13] "In the Matter of Unlicensed Operation in the TV Broadcast Bands: Second Report and Order and Memorandum Opinion and Order," Federal Communications Commission, Tech. Rep. 08-260, Nov. 2008.
- [14] Chin-Sean Sum; Villardi, G.P.; Rahman, M.A; Baykas, T.; Ha Nguyen Tran; Zhou Lan; Chen Sun; Alemseged, Y.; Junyi Wang; Chunyi Song; Chang-woo Pyo; Filin, S.; Harada, H., "Cognitive communication in TV white spaces: An overview of regulations, standards, and technology [Accepted From Open Call]," *Communications Magazine, IEEE*, vol.51, no.7, pp.138,145, July 2013.
- [15] Qiu, R.C.; Zhe Chen; Nan Guo; Yu Song; Peng Zhang; Husheng Li; Lifeng Lai, "Towards a Real-Time Cognitive Radio Network Testbed: Architecture, Hardware Platform, and Application to Smart Grid," *Networking Technologies for Software Defined Radio (SDR) Networks*, 2010 *Fifth IEEE Workshop on*, vol., no., pp.1,6, 21-21 June 2010.

# Segmentácia obrazov metódou watershed

Iveta GLADIŠOVÁ, Ján MIHALÍK,

Katedra elektroniky a multimediálnych telekomunikácií  
Fakulta elektrotechniky a informatiky, Technická univerzita Košice, SR

iveta.gladisova@tuke.sk, jan.mihalik@tuke.sk

**Abstrakt** — Článok je zameraný na segmentáciu obrazov s využitím metódy watershed. Podrobnejšie sa venuje Cannyho hranovému detektoru, ktorý sa využíval v navrhnutom simulačnom programe. Vyhodnocuje dosiahnuté výsledky pri segmentácii obrazu metódou watershed v kombinácii s použitím Cannyho hranového detektora na ukázkovom obraze, pričom je zameraný na poukázanie vplyvu nastavovania parametrov segmentácie obrazu na celkový výsledok segmentácie.

**Kľúčové slová** — segmentácia, watershed, Cannyho hranový detektor

## I. ÚVOD

V dnešnej dobe sa čoraz častejšie kladie dôraz na vizuálnu informáciu, preto je potrebné tieto informácie vhodne spracovávať. Úprava týchto informácií sa môže realizovať viacerými spôsobmi a jedným z nich je aj segmentácia obrazu. Cieľom segmentácie je jednotlivé segmenty v obraze od seba navzájom oddeliť a zároveň ich oddeliť aj od pozadia, ktoré je pre ďalšie spracovanie obrazu väčšinou nepodstatné [1], [2]. Pomocou segmentácie je možné v obraze zvýrazniť hrany alebo vybrať objekty, ktoré sú pre celkový vnem človeka najpodstatnejšie. Segmentácia je zložitý proces a pri výbere správnej segmentačnej metódy musíme brať do úvahy to, že každý obraz je iný a môže sa líšiť hlavne nejednoznačnosťou obrazových dát ako aj informačným šumom [10].

## II. SEGMENTÁCIA OBRAZU VYUŽÍVAJÚCA METÓDU WATERSHED

Prehľad metód segmentácie šedotónových obrazov spolu s výsledkami jej simulácie sme uviedli v [9]. Pri segmentácii obrazu metódou prahovania, či metódou detekcie hrán alebo metódou detekcie oblastí sa často stáva, že jednotlivé segmenty navzájom splývajú [3]. A práve oddelenie týchto segmentov zabezpečuje metóda watershed [4],[5]. Je to segmentácia založená na princípe morfolologickej topológie. Obraz je chápaný ako topografický povrch, reliéf krajiny, pozostávajúceho z údolí a pohorí. Tieto sú dané intenzitami šedotónového obrazu alebo veľkosťami gradientu. Následne segmentácia rozdelí obraz do oblastí lokálnych miním, pričom každá oblasť obsahuje len tie body, ktorých cesta najstrmšieho zostupu končí v danom minime. Za tohto predpokladu sa segmenty, teda v tomto prípade údolia, naplňajú vodou a v momente, keď hrozí, že sa voda vyleje, je tento segment ohraničený a teda aj oddelený od susedného segmentu. Transformácia watershed je založená na simulácii postupného zaplavovania kvapalinou morfologického gradientu ako topografického reliéfu. Tento krok sa realizuje pomocou vzdialenostnej funkcie.

*Vzdialenostnú funkciu* získame tzv. rezíduom, ktoré môžeme vysvetliť ako konečnú eróziu, ktorá sa na obraz aplikuje dovtedy, pokiaľ v danom okolí nezostane objekt menší, ako zvolený štruktúrálny element [11]. Problémom je, že získaním vzdialenostnej funkcie síce dostávame údolie, no v nesprávnom prevedení (otočené je naopak) a prevedením watershed na takýto obraz by sme získali nevhodný výsledok. Preto je dôležité tento obraz upraviť jednoduchou inverziou, pričom na takto upravený obraz sa už môže aplikovať watershed. Celý proces segmentácie prebieha teda v dvoch krokoch. Získanie segmentov použitím jednej z metód (prahovanie, detekcia hrán, detekcia oblastí) a oddelením získaných segmentov využitím metódy watershed.

Náš simulačný program pozostáva z dvoch častí, a to z Cannyho hranového detektora a z metódy watershed. *Cannyho hranový detektor* je považovaný za jeden z najefektívnejších postupov hľadania hrán, ktorý v sebe zahŕňa niekoľko krokov, ktorých súčasťou sú aj niektoré z metód detekcie hrán. Celý postup sa dá zhrnúť do 4 nasledujúcich krokov [7]:

1. odstraňovanie šumu,
2. výpočet gradientu,
3. výpočet lokálnych maxím,
4. prahovanie.

*Odstraňovanie šumu* - táto časť Cannyho detektora môže byť považovaná za predspracovanie obrazu na odstránenie šumu, ktorý vzniká napríklad digitalizáciou obrazu alebo nedokonalosťou snímacích zariadení. Pri tejto metóde sa najčastejšie používa Gaussov filter, pri ktorom je možnosť zvoliť si smerodajnú odchýlku a tak zmeniť hodnoty masky filtra [8]. Pri maskách filtrov (maticiach) je potrebné zvoliť jej vhodnú veľkosť so správnou smerodajnou odchýlkou pre daný konkrétny obraz. Príklad masky Gaussovho filtra o rozmere 3x3 je na Obr.1a.

$$\begin{bmatrix} 0,0113 & 0,0838 & 0,0113 \\ 0,0838 & 0,6193 & 0,0838 \\ 0,0113 & 0,0838 & 0,0113 \end{bmatrix} \quad \begin{bmatrix} 0,1111 & 0,1111 & 0,1111 \\ 0,1111 & 0,1111 & 0,1111 \\ 0,1111 & 0,1111 & 0,1111 \end{bmatrix} \quad \begin{bmatrix} 0,1667 & 0,6667 & 0,1667 \\ 0,6667 & -3,3333 & 0,6667 \\ 0,1667 & 0,6667 & 0,1667 \end{bmatrix}$$

Obr.1. Masky filtrov: a) Gaussovho, b) priemerovacieho, c) Laplaceovho.

Pri experimentoch sme používali na filtrovanie obrazu aj priemerovací filter, ktorý podobne ako Gaussov filter, je reprezentovaný maticou, ktorej rozmer môže byť ľubovoľný. Na Obr.1b je uvedený príklad priemerovacieho filtra s rozmerom bloku 3x3. Na Obr.1c je príklad matice Laplaceovho filtra veľkosti 3x3, ktorý sa používa predovšetkým na zvýraznenie hrán.

*Výpočet gradientu* – veľkosť hrany odpovedá modulu gradientu spojitej obrazovej funkcie v príslušnom obrazovom prvku (bode). Gradient ukazuje smer najväčšieho vzrastu obrazovej funkcie. V prípade hľadania hrán v obraze sa tiež využívajú konvolučné jadra, resp. konvolučné masky. Pri detekcii hrán prvou deriváciou sa najčastejšie používa Sobelova, Prewittova, Robinsonova alebo Kirschova maska [2]. Testovaním sa ukázalo sa, že najvhodnejšou je Sobelova konvolučná maska v tvaroch uvedených na Obr.2. Aplikovaním týchto dvoch masiek na originálny obraz získame obrazy reprezentujúce konvulúciu v horizontálnom a vertikálnom smere a následne sa výpočtom získa gradient.

-1	0	1
-2	0	2
-1	0	1

**X**

-1	-2	-1
0	0	0
1	2	1

**Y**

Obr.2. Sobelové konvolučné masky pre horizontálny (X) a pre vertikálny smer (Y).

*Výpočet lokálnych maxím* - úlohou tohto kroku je z gradientu získať lokálne maxima. Týmto krokom sa docieľa, že z gradientu, ktorý predstavuje obraz hrubých hrán sa získajú tenké hrany. V praxi to znamená, že sa porovnávajú jasové úrovne okolitých bodov a za bod hrany bude považovaný ten obrazový bod, ktorého jasová úroveň je väčšia ako jasová úroveň okolitých bodov. Najčastejšie sa porovnávajú jasové úrovne v horizontálnom a vertikálnom smere, keďže lokálne maximum sa počíta z gradientu, ktorý bol získaný obrazmi vertikálnej a horizontálnej konvulúcie.

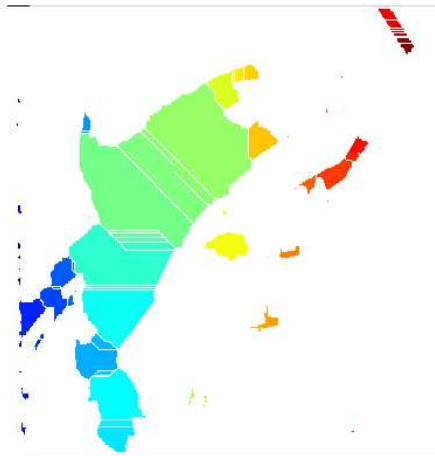
*Prahovanie* – je posledným krokom Cannyho detektora. Veľmi často sa používa prahovanie s hystereziou. To znamená, že sa zvolia 2 prahy - minimálny T1 a maximálny T2. V prípade, že jasová úroveň bodu je väčšia ako T2, je obrazový bod považovaný za bod hrany. Ak je v intervale <T1, T2>, bude bodom hrany iba v prípade, že jeho susedný bod už bol považovaný za hranu [8].

### III. EXPERIMENTÁLNE VÝSLEDKY SIMULÁCIE A ICH ZHODNOTENIE

Program na implementáciu segmentácie obrazu pomocou Cannyho detektora hrán a následným využitím metódy watershed bol realizovaný v programovom prostredí Matlab. Experimenty sme prevádzali na ôsmich vstupných farebných obrazoch rôznych veľkostí. Pre vyhodnotenie dosiahnutých výsledkov sme z nich vybrali reprezentačný obraz – lena.bmp. Vstupný farebný obraz sa pri spracovaní a reprezentácii upravoval na jeho šedý tvar (prevod farebného obrazu na šedý sa robí v programe automaticky, bez potreby zadávania ďalších parametrov).

V navrhnutom simulačnom programe pre segmentáciu obrazu s využitím metódy watershed sa najprv z obrazu odstraňuje šum pomocou dvoch predvolených filtrov (Gaussovho a Priemerovacieho). Pri Gaussovom filtri je možnosť zvoliť si smerodajnú odchýlku, pomocou ktorej dokážeme meniť hodnoty masky tohto filtra. Pri oboch typoch filtra je možnosť voľby veľkosti masky príslušného filtra, čím sa môže zabezpečiť rôzne odstránenie šumu z obrazu a následne iný výsledok segmentácie. Veľkosť okna nie je obmedzená a môže sa zvoliť aj nesúmerná veľkosť masky. Program na rozdiel od klasického Cannyho hranového detektora (iba Sobelová konvolučná maska) ponúka na výber aj 3 ďalšie konvolučné masky (Sobel, Prewitt, Roberts), pomocou ktorých sa realizuje výpočet konvolučných obrazov a na základe nich sa vypočíta gradient. Pomocou parametra segmentácie „Prah“ sa zvolí jasová úroveň prahu, čím sa zjednodušuje proces hľadania hrán v obraze. Tie body, ktorých hodnota jas je pod úrovňou prahu, sa nebudú považovať za body hrany a tým sa proces hľadania hrán urýchľuje. Nastavením parametrov pre blok „Erózia“ v programe (pri niektorých obrazoch je vhodné ju použiť) dokážeme objekty zmenšiť, resp. dokážeme odstrániť falošné hrany. Aj tu je možnosť zvoliť veľkosť bloku, s ktorým sa tento krok môže zrealizovať. Ovládacím blokom „Erózia“ sa končí proces hľadania hrán v obraze a získali sa tak segmenty v obraze, ktoré však môžu niekedy navzájom splyvať. Preto je teda vhodné tieto segmenty od seba navzájom oddeliť [6], resp. ich aj farebne reprezentovať. Na to nám poslúžila metóda watershed. Ovládacia časť simulačného programu umožňuje zvoliť: vzdialenostnú funkciu, čím sa umožňuje zvoliť približný tvar objektov v obraze, ktorý môže zabezpečiť presnejšiu segmentáciu. V tomto bloku sú na výber 4 druhy štandardných tvarov štruktúrnych elementoch a to: osemuholník, kosoštvorec, kružnica a šachovnica [11]. V procese získavania vzdialenostnej funkcie sa často môže stať, že jednotlivé segmenty aj naďalej splyvajú a dokonca sa aj prekrývajú (tým môže dôjsť k vzniku falošného segmentu). V bloku: Odstránenie prekrytia, podobne ako aj v ďalších ovládacích prvkoch programu, je možnosť voľby veľkosti bloku, s ktorým sa pracuje. Týmto blokom sa vlastne zabezpečuje jemnejšia segmentácia obrazu.

Na vybranom ukázkovom obraze „lena.bmp“ si ukážeme vplyv voľby jednotlivých parametrov a ich nastavenie tak, aby sme získali čo najlepšie výsledky segmentácie. Na Obr. 3 je zobrazený originálny obraz a jeho segmentovaný tvar pri nastavení uvedených parametrov. Môžeme si všimnúť, že proces segmentácie nie je dokonalý. Je to spôsobené zadanými hodnotami ovládacích prvkov pre segmentáciu, ktoré nie sú pre tento obraz vhodne nastavené.



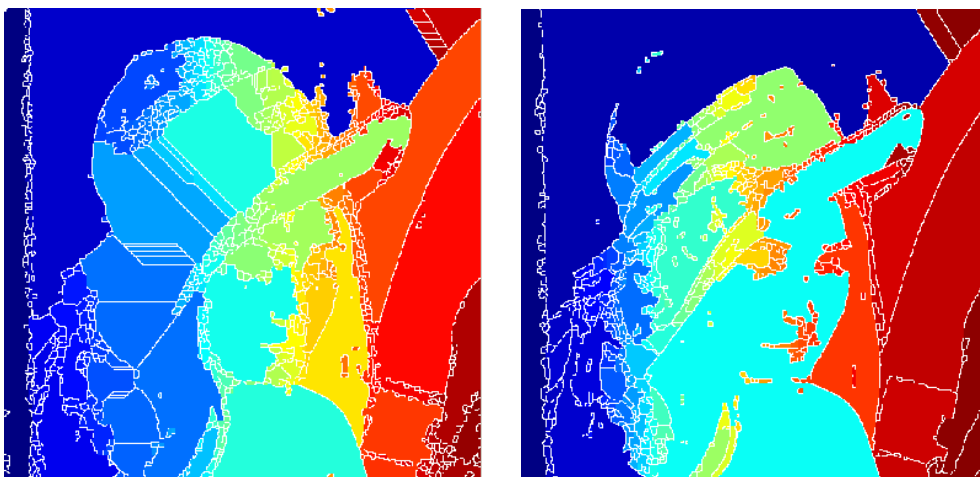
Obr.3. Vľavo: originálny obraz „lena.bmp“; vpravo: jeho segmentovaný obraz s parametrami: Gaussov filter s rozmerom bloku 3x3 a smerodajnou odchýlkou 0,7; Konvolučná maska: Sobel, prah: 0,4; Erózia: rozmer bloku 3x3, Vzdialenostná funkcia: 8-uholník; odstránenie prekrytia: rozmer bloku 3x3.

Problém nedostatočnej segmentácie môže byť spôsobený napríklad aj eróziou, ktorá odstraňuje segmenty, ktorých rozmer je menší ako 3x3. Pre tento obraz je vhodné tento krok vynechať, čo znamená, že žiadne segmenty nebudú odstraňované a segmentácia tak bude jemnejšia. Vynechaním tohto kroku sme získali presnejší výsledok segmentácie, ako môžeme vidieť na Obr.4 vľavo. Ďalším problémom môže byť aj nevhodne zvolená veľkosť filtra v procese predspracovania. Pri zmene parametrov Gaussovho filtra na veľkosť masky 7x7 a so smerodajnou odchýlkou 1,4 sme dosiahli ešte lepší výsledok (viď. Obr. 4 vpravo).

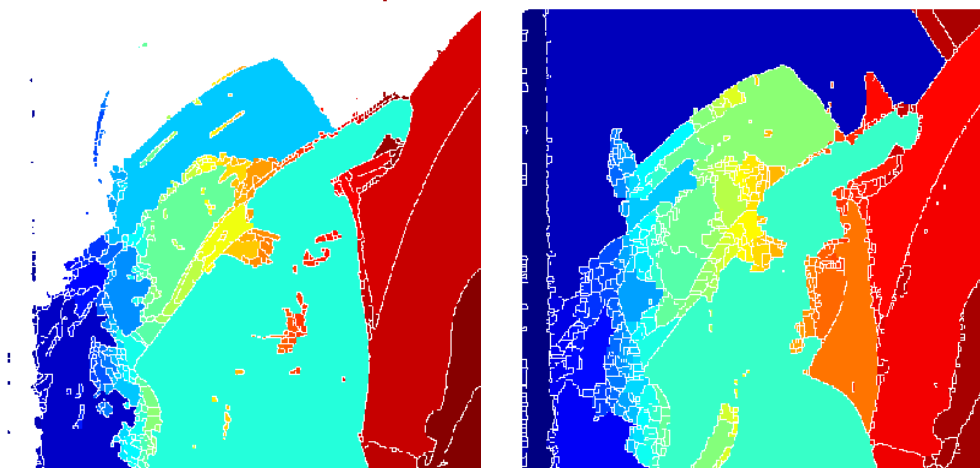
Tento výsledok segmentácie podáva celkom dobré informácie o objektoch v obraze. Avšak skúsili sme zmeniť aj ďalšie parametre programu, napríklad iný typ konvolučnej masky. Výberom Robertsovej konvolučnej masky sme získali nasledujúci segmentovaný obraz (Obr. 5 vľavo). Môžeme konštatovať, že je horší ako predchádzajúci obraz (Obr.4 vpravo), pretože tvár



nám splynula s rukou. Tento obraz sme sa pokúsili opraviť pomocou bloku Odstránenia prekrytia, keď sme zvolili veľkosť bloku 7x7 (viď Obr.5 vpravo).



Obr.4. ..Vľavo: Segmentovaný obraz Lena: Gaussov filter s rozmerom bloku 3x3, smerodajnou odchýlkou 0,7; Konvolučná maska: Sobel, prah: 0,4; Erózia: žiadna; Vzdialenostná funkcia: 8-uholník; Odstránenie prekrytia: rozmer bloku 3x3; vpravo: segmentovaný obraz Lena: Gaussov filter s rozmerom bloku 7x7, smerodajnou odchýlkou 1,4; vynechané Odstránenie prekrytia; ostatné parametre sú také isté ako v prípade a).

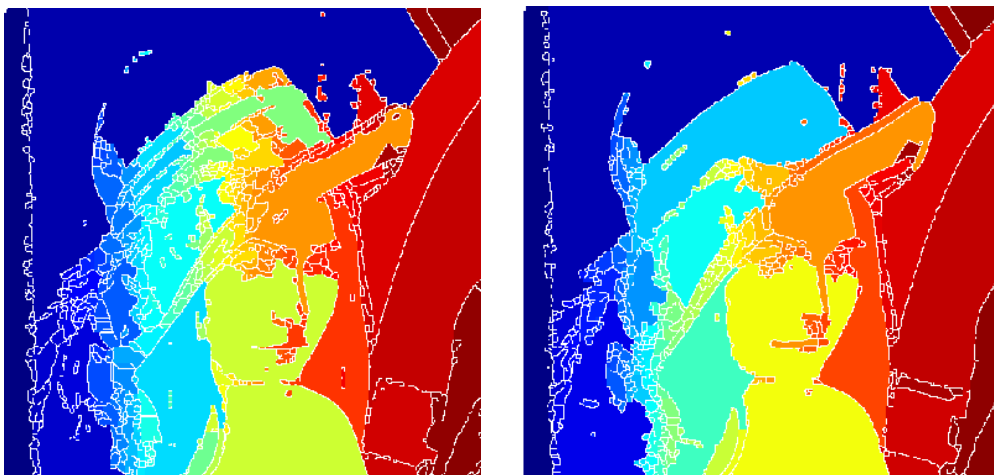


Obr. 5. Vľavo: Segmentovaný obraz Lena: Gaussov filter s rozmerom bloku 7x7, smerodajnou odchýlkou 1,4; Konvolučná maska: Roberts, prah: 0,4; Erózia: žiadna; Vzdialenostná funkcia: 8-uholník; Odstránenie prekrytia: rozmer bloku 3x3; vpravo: Segmentovaný obraz Lena: oproti predchádzajúcim parametrom došlo k zmene iba pri parametri: Odstránenie prekrytia: rozmer bloku 7x7.

Úpravou sme síce získali lepší výsledok ako na Obr.5 vľavo, avšak nezvýraznili sa nám časti tváre (oči, ústa, nos) ako Obr.4 vpravo (v niektorých prípadoch však takýto výsledok segmentácie môže postačovať). Obr. 4 vpravo sme považovali zatiaľ za najlepší výsledok segmentácie a pokúsili sme sa ho ešte vylepšiť zmenou prahu konvolučnej masky na hodnotu 0,35 (Obr.6 vľavo). Zmenou prahu sme získali výsledok, ktorý môžeme považovať za podstatne lepší, ako segmentáciou na Obr.4 vpravo.

Ďalší vhodný výsledok segmentácie sme získali zámennou Gaussovho filtra za priemerovací filter a zámennou Sobelovej konvolučnej masky za Prewittovu masku (Obr.6 vpravo). V tomto prípade sa zvýraznili ústa a nos, čiže týmito parametrami sme získali viac informácií o tvári, avšak menej informácií o klobúku. Obraz môžeme jemnejšie segmentovať voľbou veľkosti bloku odstránenia prekrytia, čím získame detailnejšie informácie o obraze (nie je to nutné). Z globálneho hľadiska najviac informácií o celom obraze podáva obraz na Obr.6 vľavo. V tomto obraze sú zjavné detaily tváre, ale aj ostatných častí obrazu. No pri samotnej segmentácii si musíme uvedomiť, na čo sa chceme zamerať a tomu prispôbiť voľbu parametrov, aby výsledok segmentácie nám poskytol dostatok informácií.





Obr.6. Vľavo: Segmentovaný obraz Lena: Gaussov filter s rozmerom bloku 7x7, smerodajnou odchýlkou 1,4; Konvolučná maska: Sobel, prah: 0,35; Erózia: rozmer žiadna, Vzdialenostná funkcia: 8-uholník, Odstránenie prekrytia: rozmer bloku 3x3; vpravo: Segmentovaný obraz Lena: Priemerovací filter s rozmerom 7x7, Konvolučná maska: Prewitt, prah: 0,4; Erózia, Vzdialenostná funkcia a Odstránenie prekrytia zostávajú rovnaké ako v prípade a).

#### IV. ZÁVER

Zo všeobecného hľadiska neexistuje univerzálna metóda segmentácie, ktorá by sa dala použiť na spracovanie každého obrazu. Pri simulácii segmentácie obrazu metódou watershed môžeme konštatovať, že aj konkrétna metóda môže poskytovať rôzne výsledky podľa toho, aké parametre bude využívať. Každý obraz obsahuje šum a pri snahe odstrániť tento nežiaduci efekt boli použité rôzne filtre s odlišnými hodnotami, čím sme pri celkovej segmentácii získali rôzne výsledky. Ďalším poznatkom je, že na spracovanie obrazu má výrazný vplyv aj výber správnej konvolučnej masky. Takýchto masiek je niekoľko a pri segmentácii týmito maskami sa dosahujú opäť odlišné výsledky. Výrazný vplyv na celkovú segmentáciu má aj hodnota prahu. Tento parameter sa využíva z dôvodu urýchlenia hľadania hrán v obraze. Znižovaním prahovej hodnoty zvyšujeme počet detailov v obraze a v konečnom dôsledku sa zvyšuje aj počet segmentov v samotnom spracovávanom obraze. V niektorých prípadoch je práve väčší prah výhodou (ak nevyžadujeme vysoký počet detailov v obraze), no v iných je nevýhodou, pretože strácame informácie o obraze, čo môže skomplikovať zhodnotenie informácie.

Na záver by sme mohli konštatovať, že segmentácia je zložitý proces, v ktorom je potrebné vybrať správnu metódu a následne pri tejto metóde navoliť správne parametre, aby výsledok poskytoval dostatok pre nás potrebných a postačujúcich informácií. Jednotlivé kroky pri segmentácii medzi sebou úzko súvisia a tým ovplyvňujú celkový proces spracovania.

#### LITERATÚRA

- [1] Grady, L.: *Random Walks for Image Segmentation*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 11, 2006, pp 1-17.
- [2] Pal, S.K. – Pal, N.R., *A Review on Image Segmentation Techniques*, Pattern Recognition Letters, vol. 26, no. 9, 1993.
- [3] Hoover, A. – Jean-Baptiste, G., *An Experimental Comparison of Range Image Segmentation Algorithms*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 7, July 1996.
- [4] Eddins, S., *The Watershed Transform: Strategies for Image Segmentation*, from online writing: <http://www.mathworks.com/company/newsletters/articles/the-watershed-transform-strategies-for-image-segmentation.html>
- [5] Belaid, L.J. - Mourou, W., *Image segmentation: a watershed transformation algorithm*, Image Anal Stereol 2009; 28:93-102, from online writing: <http://www.ias-iss.org/ojs/IAS/article/viewFile/852/755>.
- [6] Amine, K.- Farida, M.H., *Range image segmentation classification*, IJCSI International Journal of Computer Science Issues, vol. 9, Issue 4, No 1, July 2012.
- [7] Canny, J. F., *A Computational Approach to Edge Detection*, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, 1986.
- [8] Law, T. – Itoh, H. – Seki, H.: *Image Filtering, Edge Detection and Edge Tracing Using Fuzzy Reasoning*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 5, 1996.
- [9] Mihalík, J. - Gladišová, I.: *Implementácia metód segmentácie obrazov*. In: Electrical Engineering and Informatics 5 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, FEI TU, 2014.
- [10] Shi, J. - Malik, J.: *Normalized Cuts and Image Segmentation*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, 2000, pp 888-905.
- [11] Gladišová, I. - Mihalík, J., *Aplikácia vybraných morfológických transformácií na spracovanie obrazu*. In: Electrical Engineering and Informatics 4, Proc. of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, 2013, p. 541 - 546.

# Implementácia metód segmentácie obrazov

Ján MIHALÍK, Iveta GLADIŠOVÁ,

Katedra elektroniky a multimediálnych telekomunikácií  
Fakulta elektrotechniky a informatiky, Technická univerzita Košice, SR

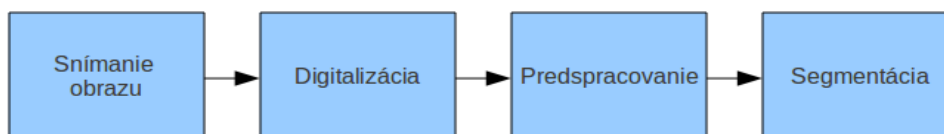
jan.mihalik@tuke.sk, iveta.gladisova@tuke.sk,

**Abstrakt** — Článok sa zaoberá segmentáciou statických obrazov, ktorej cieľom je nájsť segmenty v obraze. V stručnosti popisuje známe segmentačné metódy - segmentáciu prahovaním, segmentáciu pomocou detekcie hrán, segmentáciu pomocou detekcie oblastí a segmentáciu s využitím metódy watershed. V užívateľskom programovom prostredí Matlab bol vypracovaný simulačný program realizujúci segmentáciu obrazu. V článku sú vyhodnotené dosiahnuté výsledky, pričom sa zameriava na poukázanie vplyvu nastavovania parametrov segmentácie obrazu na celkový výsledok segmentácie.

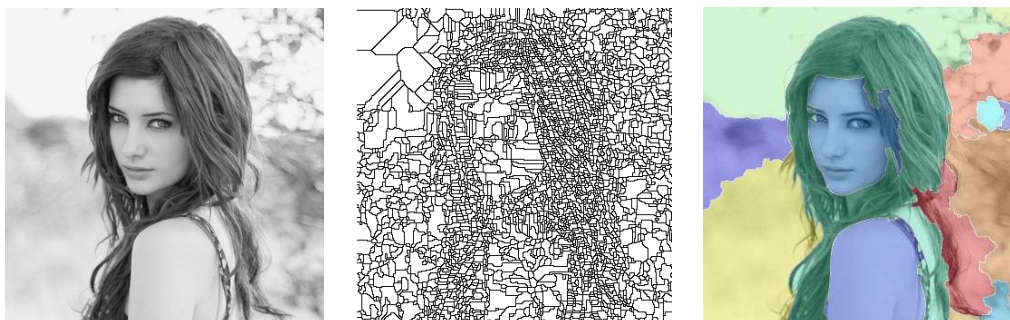
**Kľúčové slová** — segmentácia hrán, segmentácia oblastí, konvolučné masky, gradient

## I. ÚVOD

Úlohou segmentácie je rozdeliť spracovávaný obraz na menšie časti tzv. segmenty, ktoré majú silnú koreláciu. Pri segmentácii je snaha jednotlivé segmenty od seba navzájom oddeliť a zároveň ich oddeliť aj od pozadia, ktoré je pre ďalšie spracovanie obrazu väčšinou nepodstatné [1], [5]. Ide vlastne o analýzu obrazu, ktorej cieľom je obraz popísať alebo ho zaradiť do určitej triedy, teda klasifikovať. Segmentáciou sa jednotlivé časti obrazu identifikujú podľa vlastností, ktoré sú pre ne spoločné a na základe toho sa v obraze vytvorí segment. Segmentácia obrazu prebieha v niekoľkých krokoch, ktoré popisuje nasledujúci obr.1. Aby sme mohli obraz segmentovať, musíme ho najprv získať pomocou snímacích zariadení (napr. kamera, skener, fotoaparát, atď.). Niekedy sa môže stať, že získaný obraz je v analógovej forme. Preto je potrebné tento obraz previesť do digitálnej formy pomocou A/D prevodníkov. Pred samotnou segmentáciou je potrebné obraz ešte upraviť, zbaviť ho nadbytočnej informácie (šumu), resp. zvýrazniť hrany. Tento proces sa nazýva predspracovanie. Samotný proces segmentácie sa aplikuje na takto upravený obraz využitím jednej z metód segmentácie, ktoré budú popísané nižšie. Na obr.2 je zobrazená segmentácia konkrétneho obrazu.



Obr.1. Segmentácia obrazu.



Obr. 2 Ukážka pôvodného obrazu (vľavo), segmentovaného (stred) a segmentovaného s predspracovaním (vpravo).

## II. METÓDY SEGMENTÁCIE

Segmentácia je zložitý proces, ktorý sa uplatňuje na rôzne typy obrazu. Pri výbere správnej segmentačnej metódy musíme brať do úvahy to, že každý obraz je iný, výnimočný a môže sa líšiť hlavne: nejednoznačnosťou obrazových dát a informačným šumom. Z toho dôvodu existujú rôzne druhy segmentácie, ktoré sa vyznačujú svojimi charakteristickými vlastnosťami a k segmentácií pristupujú odlišnými spôsobmi. Na základe toho môžeme segmentáciu šedotónových obrazov [1], [2] rozdeliť na: segmentáciu prahovaním, segmentáciu pomocou detekcie hrán a segmentáciu pomocou detekcie oblastí. V stručnosti uvedieme jednotlivé segmentácie.

**Prahovanie** je najjednoduchší spôsob detekcie celých oblastí v obraze. Patrí medzi metódy, ktoré sú založené na obrazových bodoch a v nekomplikovaných prípadoch ide o veľmi rýchlu a spoľahlivú metódu [5], [7]. Pri prahovaní sa vychádza z predpokladu, že každý obrazový prvok (bod) má určitú jasovú úroveň. Následne sa zvolí určitá hodnota jasu, resp. interval jasových úrovní a všetky obrazové prvky, ktorých jasová úroveň je väčšia ako zvolený prah, resp. sú v rozsahu zvoleného intervalu, vytvoria oblasť segmentu. Výsledkom prahovania je binárny obraz. Prahovacie metódy môžeme rozdeliť na: jednoduché a adaptívne. Jednoduché prahovanie poznáme: *globálne* (zvolí sa jedná úroveň prahu pre celý obraz), *poloprahovanie* (obrazové prvky s jasovou úrovňou menšou, ako zvolený prah, sa stávajú pozadím a tie, s hodnotou väčšou ako zvolený prah, si ponechávajú svoju jasovú hodnotu), *spektrálne* (nevolí sa iba jeden prah, ale za prah sa zvolí interval jasových úrovní) a *multispektrálne* (je to obdoba spektrálneho prahovania, no v tomto prípade sa volí viacej intervalov). Adaptívne prahovanie poznáme: *Chow\_Kaneno* (obraz sa rozdelí na menšie časti, z histogramu jednotlivých častí sa pre ne určí optimálny prah a prahová hodnota pre konkrétny bod sa vypočíta interpoláciou prahových hodnôt jednotlivých častí) a *lokálne* (princíp tejto metódy spočíva v tom, že prahovú hodnotu pre konkrétny bod vypočítame z malého okolia tohto bodu ako priemer jasových hodnôt v okolí, alebo mediánom jasových úrovní bodov okolia, alebo strednou hodnotou minimálnej a maximálnej jasovej hodnoty v okolí bodu).

**Segmentácia pomocou detekcie hrán** je metóda, pri ktorej sa získavajú obrazové body, v ktorých sa výrazne mení jas. Princíp tejto metódy pozostáva z dvoch krokov. Prvým krokom je nájdenie hrán v obraze a druhým krokom je vyplnenie oblastí, ktoré vznikli uzavretím jednotlivých hrán [4]. Hranový bod je taký bod v obraze, ktorý sa vyznačuje veľkým modulom gradientu obrazovej funkcie. Gradientná zmena môže prebiehať rôzne, čo môže byť spôsobené rýchlosťou alebo priebehom tejto zmeny. Detekcia hrán v obraze sa môže realizovať viacerými postupmi, medzi ktoré patrí aj: *detekcia hrán prvou deriváciou*, *detekcia hrán druhou deriváciou*, *Cannyho detektor* a *Houghova metóda*.

**Segmentácia pomocou oblastí** - princíp týchto metód spočíva v tom, že na rozdiel od metódy segmentácie pomocou detekcie hrán sa pri tejto metóde priamo získavajú segmenty [3]. Na získanie segmentov v obraze sa musí zvoliť určité kritérium, pomocou ktorého budú jednotlivé segmenty vznikať. Medzi takéto kritérium patrí hlavne farba, jas, textúra a pod. Takúto segmentáciu môžeme rozdeliť podľa toho, ako budú segmenty vznikať, na: *spájanie oblastí*, *delenie oblastí* a *na delenie a spájanie oblastí súčasne*. Už podľa názvu metód je zjavné, o čo v týchto metódach pôjde. Výhodou je väčšia odolnosť voči šumu a vzhľadom k ostatným metódam segmentácie dávajú lepšie výsledky. Problém je však v tom, že jednotlivými metódami môžeme dostať rôzne výsledky spôsobené rôznym začiatkom spracovania obrazu. Výsledok segmentácie so začiatkom spracovania v ľavom hornom rohu obrazu sa môže líšiť od výsledku segmentácie so začiatkom spracovania v pravom dolnom rohu pri tej istej segmentácii. Ďalšie metódy segmentácie šedotónových obrazov môžeme nájsť v [6].

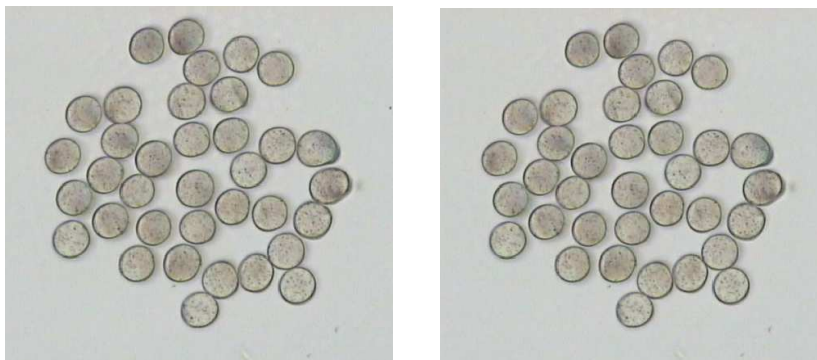
**Segmentácia pomocou watershed** je založená na princípe morfolologickej topológie. Obraz je chápaný ako topografický povrch, reliéf krajiny, pozostávajúceho z údolí a pohorí [10]. Tieto sú dané intenzitami šedotónového obrazu alebo veľkosťami gradientu. Následne segmentácia rozdelí obraz do oblastí lokálnych miním, pričom každá oblasť obsahuje len tie body, ktorých cesta najstrmšieho zostupu končí v danom minime. Za tohto predpokladu sa segmenty, teda v tomto prípade údolia, naplňujú vodou a v momente, keď hrozí, že sa voda vyleje, je tento segment ohraničený a teda aj oddelený od susedného segmentu.

## III. EXPERIMENTÁLNE VÝSLEDKY SIMULÁCIE

Program na implementáciu vybraných algoritmov segmentácie bol vytvorený v užívateľskom programovom prostredí Matlab. Umožňuje prácu s predefinovanými obrazmi (binárnymi šedotónovými, aj farebnými), ktoré boli vybrané tak, aby čo najlepšie demonštrovali možnosti segmentácie obrazu. Program umožňuje načítanie aj iných, ako predvolených obrazov typu jpeg, bmp, gif. Filtrovanie načítaného obrazu odstraňuje šum pomocou dvoch predvolených filtrov

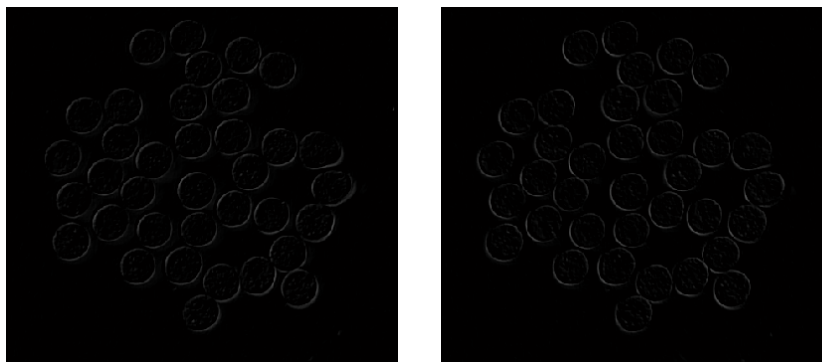
(Gaussovho a priemerovacieho) [8]. Pri Gaussovom filtri je možnosť zvoliť si smerodajnú odchýlku, pomocou ktorej dokážeme meniť hodnoty masky filtra. Pri oboch typoch filtra je možnosť voľby veľkosti masky príslušného filtra, čím sa môže zabezpečiť rôzne odstránenie šumu z obrazu a následne iný výsledok segmentácie. Veľkosť okna nie je obmedzená, čo znamená, že sa môže zvoliť aj nesúmerná veľkosť napríklad 3x2, kde 3 predstavuje počet riadkov a 2 počet stĺpcov. Na obr.3 môžeme vidieť vplyv veľkosti bloku Gaussovho filtra pri odstránení šumu obrazu „mince.bmp“.

Program má na výber 3 konvolučné masky (Sobel, Prewitt, Roberts), pomocou ktorých sa realizuje výpočet konvolučných obrazov, na základe ktorých sa vypočítava gradient. Gradient je fyzikálna veličina a vyjadruje vektor prvých derivácií funkcie podľa jednotlivých premenných. Cieľom gradientu je určiť smer a veľkosť najväčšej zmeny tejto funkcie. Samotná detekcia hrán prvou deriváciou využíva výpočet konvolúcie obrazu so zvolenou konvolučnou maskou.



Obr.3. Filtrovaný obraz Gaussovým filtrom s rozmerom bloku 2x2 (vľavo) a s rozmerom bloku 5x5 (vpravo), pričom v oboch prípadoch bola veľkosť smerodajnej odchýlky 0,7.

Konvolúcia je proces, kedy konvolučnú masku preložíme na príslušné miesto originálneho obrazu. Následne vynásobíme hodnoty obrazových bodov skúmaného obrazu s im zodpovedajúcimi hodnotami konvolučnej masky a urobíme súčet všetkých týchto hodnôt. Tým získame jeden nový obrazový bod reprezentujúci bod hrany v horizontálnom resp. vertikálnom smere. Tento proces sa aplikuje na každý bod obrazu, čím sa získa príslušný obraz konvolúcie v danom smere. Na obr.4 a obr.5 máme možnosť porovnať vplyv výberu konvolučnej masky na výsledný konvolučný obraz v smeroch osí x a osí y.

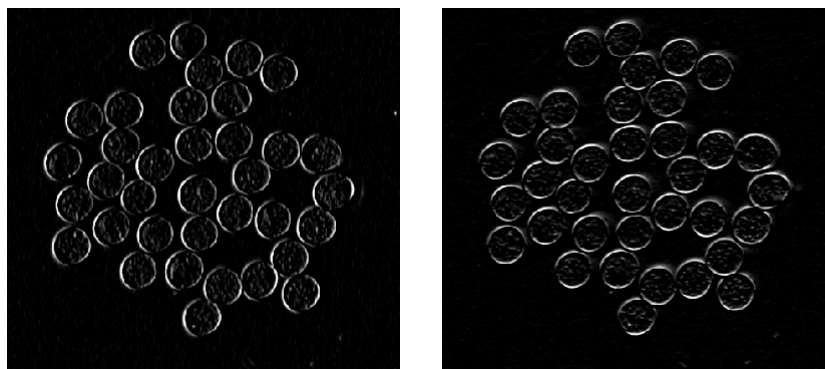


Obr.4. Konvolúcia Robertsovou maskou v smere osi X (vľavo) a v smere osi Y (vpravo).

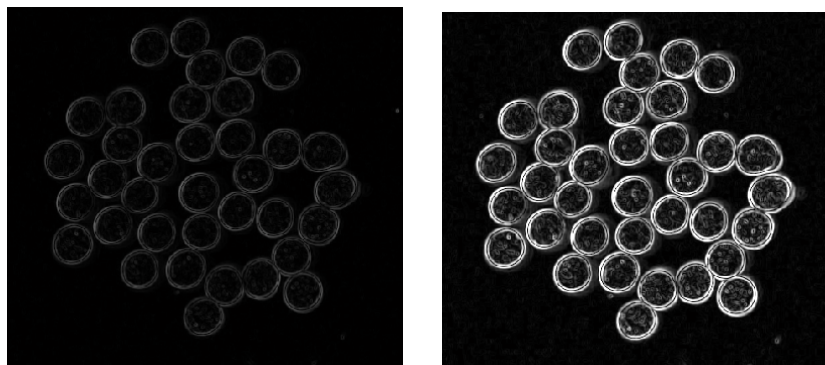
Použitím rôznych masiek získavame aj rôzne výsledky, z čoho je možné vyvodit' záver, že neexistuje univerzálna maska pre každý obraz. Práve preto je dôležité vybrať vhodnú masku, pre dosiahnutie čo najlepšieho výsledku. Hlavnou nevýhodou tejto metódy je, že po získaní gradientu síce získame hrany, ale tieto hrany sú reprezentované hrubými čiarami, ktoré by v ideálnom prípade mali byť tenké. Na obr. 6 máme možnosť vidieť gradienty odpovedajúcich obrazov a na obr.7 môžeme vidieť získané hrany týchto obrazov.

Po skončení procesu hľadania hrán v obraze (Cannyho hranový detektor) sme použitím metódy watershed [9] získali požadované segmenty v obraze (obr.8), ktoré však môžu niekedy navzájom splývať. Preto je teda vhodné tieto segmenty navzájom od seba oddeliť, resp. ich aj farebne reprezentovať (obr.9).

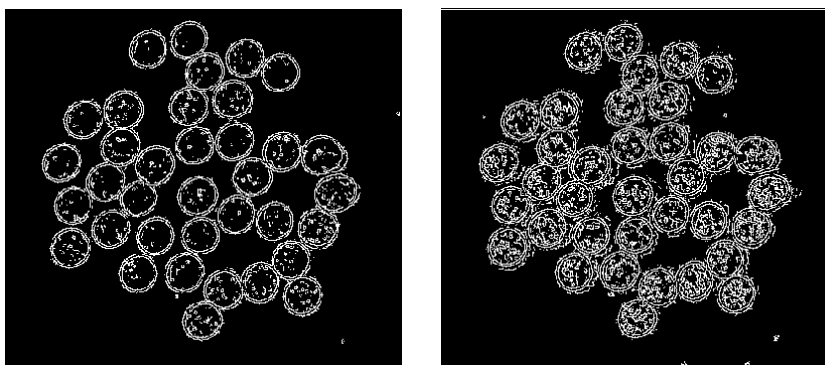




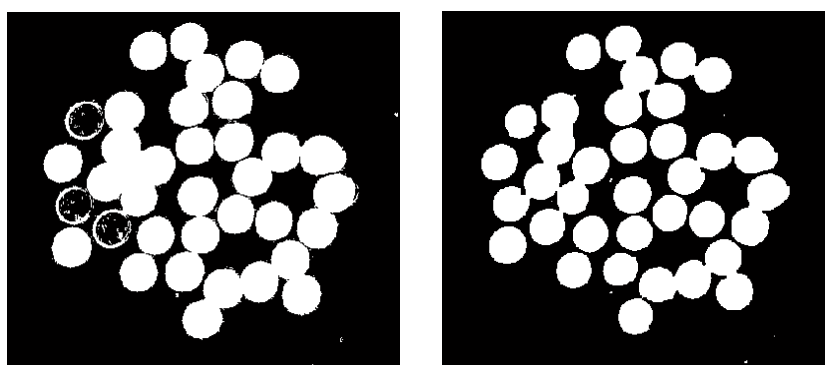
Obr.5. Konvolúcia Sobelovou maskou v smere osi X (vľavo) a v smere osi Y (vpravo).



Obr.6. Gradient po konvolúcii Robertsovou maskou (vľavo) a Sobelovou maskou (vpravo).

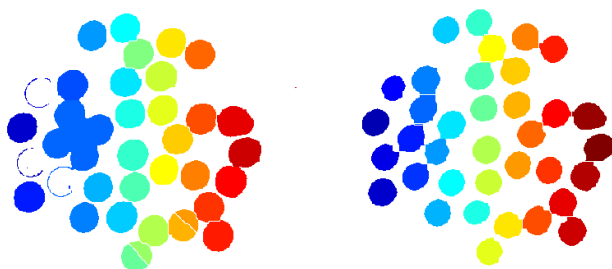


Obr.7. Hrany po konvolúcii Robertsovou maskou (vľavo) a Sobelovou maskou (vpravo).



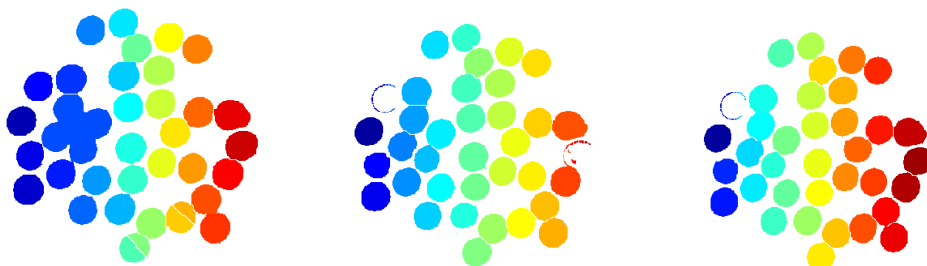
Obr.8. Segmenty po konvolúcii Robertsovou maskou (vľavo) a Sobelovou maskou (vpravo).

Z obr.9 vľavo je vidieť, že výsledok nespĺňa predpoklady správnej segmentácie. Niektoré segmenty nevznikli, čo mohlo byť spôsobené napríklad zlou voľbou prahu, iné segmenty boli rozdelené na menšie časti. Problém spôsobený neuzavretím všetkých segmentov sme sa pokúsili odstrániť nastavením hodnoty prahu konvolučnej masky na 0,4 a takto získaný segmentovaný obraz je na obr.10 vľavo. Zvolením menšieho prahu sme docielili, že sa segmenty uzavreli. Ďalším problémom v tomto segmentovanom obraze je, že niektoré segmenty splynuli do jedného segmentu.



Obr.9. Farebné rozlíšenie segmentov po konvolúcii Robertsovou maskou (vľavo) a Sobelovou maskou (vpravo).

Tento problém sme skúsili vyriešiť voľbou priemerovacieho filtra s veľkosťou masky 3x3 a zmenou štruktúrneho elementu na kosoštvorec. Tým sme získali segmentáciu obrazu, ktorú môžete vidieť na obr.10 v strede. Zadanými parametrami sme zabezpečili oddelenie jednotlivých segmentov navzájom od seba a dokonca aj spojenie dvoch častí jedného segmentu, no niektoré segmenty nám v tomto prípade neuzavrelo. Zmenili sme typ filtra na Gaussov veľkosti 3x3 a so smerodajnou odchýlkou 1,4 (viď obr.10 vpravo).



Obr.10. (Vľavo) je segmentovaný obraz s parametrami: Gaussov filter s rozmerom 2x2 a smerodajnou odchýlkou 0,7; Konvolučná maska: Roberts, prah: 0,4; (stred) segmentovaný obraz s parametrami: Priemerovací filter s rozmerom 3x3; Konvolučná maska: Roberts, prah: 0,4; (vpravo) segmentovaný obraz s parametrami: Gaussov filter s rozmerom 3x3 a smerodajnou odchýlkou 1,4; Konvolučná maska: Roberts, prah: 0,4.

#### IV. ZÁVER

Cieľom článku bolo poukázať na metódy segmentácie statického obrazu a vykonať ich implementáciu. Táto oblasť spracovania obrazovej informácie je pomerne rozsiahla. Zo všeobecného hľadiska neexistuje univerzálna metóda segmentácie, ktorá by sa dala použiť na spracovanie každého obrazu. Segmentácia je zložitý proces, v ktorom je potrebné vybrať správnu metódu a následne pri tejto metóde nastaviť správne parametre, aby výsledok poskytoval dostatok potrebných a pre nás postačujúcich informácií. Jednotlivé kroky pri segmentácií medzi sebou úzko súvisia a tým ovplyvňujú celkový proces spracovania obrazu.

#### LITERATÚRA

- [1] Pal, S.K. – Pal, N.R., *A Review on Image Segmentation Techniques*, Pattern Recognition Letters, vol. 26, no. 9, 1993.
- [2] Hoover, A. – Jean-Baptiste, G., *An Experimental Comparison of Range Image Segmentation Algorithms*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 7, July 1996
- [3] Adams, R. – Bischof, L., *Seeded Region Growing*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 6, June 1994.
- [4] Revol, Ch. – Jourlin, M., *A New Minimum Variance Region Growing Algorithm for Image Segmentation*, Pattern Recognition Letters 18, 1997.
- [5] Hlaváč V., Sedláček M., *Zpracování signálů a obrazů*. Vydavatelství ČVUT, Praha, 2000.
- [6] Hoffman, R. – Jain, A., *Segmentation and Classification of Range Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, no. 5, 1987.
- [7] Perez, A. – Gonzalez, R., *An Iterative Thresholding Algorithm for Image Segmentation*, IEEE Pattern Analysis Machine Intelligence 9, 1987.
- [8] Law, T. – Itoh, H. – Seki, H.: *Image Filtering, Edge Detection and Edge Tracing Using Fuzzy Reasoning*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 5, 1996.
- [9] Gladišová, I. – Mihalík, J.: *Segmentácia obrazov metódou watershed*. In: Electrical Engineering and Informatics 5 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, FEI TU, 2014.
- [10] Belaid, L.J. - Mourou, W., *Image segmentation: a watershed transformation algorithm*, Image Anal Stereol 2009; 28:93-102, from online writing: <http://www.ias-iss.org/ojs/IAS/article/viewFile/852/755>.

# Kalibrácia ohniskových vzdialeností kamery

<sup>1</sup>Ondrej KOVÁČ, <sup>2</sup>Ján MIHALÍK

<sup>1</sup> Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

<sup>1</sup>Ondrej.Kovac@tuke.sk, <sup>2</sup>Jan.Mihalik@tuke.sk

**Abstrakt** — Príspevok je venovaný teoretickému rozboru problému kalibrácie kamery. V praxi sa používajú tri prístupy ku kalibrácii. Prvým prístupom je kalibrácia na základe korešpondencie bodu v 3R priestore a v priestore snímky. Druhou metódou je, ak ku kalibrácii postačuje použitie 2R kalibračnej roviny a teda kalibráciu je možné vykonať aj bez presného určenia hĺbkovej súradnice korešpondujúcich bodov v 3R priestore. Najmodernejšou technikou je takzvaná autokalibrácia, ku ktorej nie je potrebné snímať kalibračný obrazec. V našom príspevku popíšeme zjednodušenú metódu kalibrácie ohniskových vzdialeností v zmysle minimalizácie strednej kvadratickej chyby za použitia 2R roviny umiestnenej v 3R vizuálnej scéne. Popíšeme exaktný postup získania ohniskových vzdialeností kamerového systému na základe jednej korešpondencie ako aj techniku založenú na viacerých korešpondenciách.

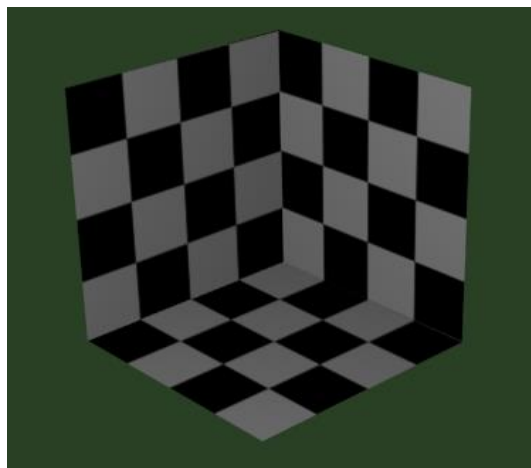
**KLúčové slová** — Kalibrácia kamery, ohnisková vzdialenosť, perspektívna projekcia

## I. ÚVOD

Kalibrácia kamerového systému (KS) je nevyhnutným krokom v 3R počítačovej grafike za účelom extrakcie informácie o rozmeroch z 2R snímok. Znalosť vonkajších, extrinzičných a vnútorných, intrinzičných parametrov je nevyhnutná pre aplikáciu rekonštrukcie 3R vizuálnej scény (VS) pomocou stereoskopického KS (SKS). Kalibrácia sa taktiež využíva pri tvorbe panoramatických fotografií kde je nevyhnutné odstrániť vplyvy rotácie a prípadného posunutia KS voči snímanej 3RVS. Pomocou kalibrácie je možné zistiť, aké optické deformácie a nepresnosti KS pri snímaní 3RVS vnaša do výslednej snímky a tieto chyby optiky kompenzovať. Techniky kalibrácie je možné rozdeliť do dvoch základných skupín [1]. Sú to tradičné techniky, ktoré využívajú kalibračný obrazec a moderné autokalibračné techniky. V tradičných metódach využívajúcich kalibračný obrazec je presnosť kalibrácie závislá od presnosti merania kalibračného obrazcu, teda je nutná presná extrakcia súradníc bodov kalibračného obrazcu [2]. Tradičné techniky je možné rozdeliť na také, ktoré využívajú korešpondenciu súradníc bodov v 3R priestore a im zodpovedajúcich bodov v rovine snímky a techniky, pri ktorých sa využíva korešpondencia bodov umiestnených v rovine umiestnenej v 3R priestore [3]. Kalibráciou je možné získať ohniskové vzdialenosti, vektor posunutia s rotáciou KS. V nami navrhovanej metóde popíšeme techniku kalibrácie ohniskových vzdialeností KS.

## II. MATEMATICKÝ MODEL KAMERY

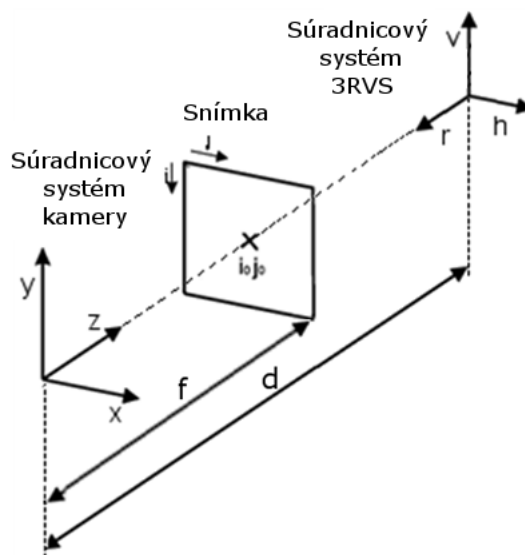
Kalibrácia je docieľaná pozorovaním kalibračného objektu ktorého geometrické parametre v 3R priestore sú s vysokou presnosťou známe. Kalibráciu je možné dosiahnuť veľmi efektívne [4]. Kalibračný objekt zvyčajne pozostáva z dvoch alebo troch ortogonálnych rovín. Niekedy je využitá iba jedna rovina, ktorá sa posúva v hĺbkovej súradnici [5] čím je možné dosiahnuť 3R referenciu. Korešpondenčné body v 3R rovine môžu byť reprezentované pomocou ortogonálnych šachovnic zobrazených na obr. 1.



Obr. 1 Kalibračná pomôcka s tromi ortogonálnymi šachovnicami.

### A. Perspektívna projekcia

Na obr.2 je zobrazené zjednodušené usporiadanie súradnicového systému kamery a 3RVS. Fyzickú kameru je možné popísať matematickým modelom perspektívnej projekcie. Pri tejto projekcii sú všetky body 3RVS premietnuté do roviny snímky cez ohnisko systému.



Obr. 2 Zjednodušené usporiadanie súradnicového systému kamery, 3RVS a snímky.

Pred výpočtom priemetu bodu 3RVS do roviny snímky je potrebná transformácia [6] súradnicového systému 3RVS do súradnicového systému kamery takto

$$\begin{bmatrix} y \\ v \end{bmatrix} = \begin{bmatrix} h \\ d - r \end{bmatrix} \quad (1)$$

V ďalšej časti budeme uvažovať bod v súradnicovom systéme kamery. Nech bod  $\mathbf{M}$  má súradnicu  $[x, y, z]$  a jeho priemet v roviny snímky je  $\mathbf{m} = [j, i]$ . Vzťah medzi súradnicou bodu  $\mathbf{M}$  v 3R priestore a jeho prietom v roviny snímky je:

$$\mathbf{m} = \mathbf{A}[\mathbf{R} \quad \vec{\mathbf{t}}]\mathbf{M} = \mathbf{P}\mathbf{M} \quad (2)$$

kde  $\mathbf{A}$  je matica vnútorných parametrov kamery

$$\mathbf{A} = \begin{bmatrix} f_x & 0 & j_0 \\ 0 & -f_y & i_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$



kde  $f_x$  a  $f_y$  sú ohniskové vzdialenosti a  $i_0$  a  $j_0$  je bod predstavujúci stred snímky a matica  $[\mathbf{R} \ \bar{\mathbf{t}}]$  sú vonkajšie parametre kamery, reprezentujúce rotáciu a posunutie kamerového systému.

$$[\mathbf{R} \ \bar{\mathbf{t}}] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \quad (4)$$

Potom matica  $\mathbf{P}$  bude projekčná matica kamery.

### III. KALIBRÁCIA POMOCOU KOREŠPONDENCIE BODOV V 3R PRIESTORE

Klasické metódy sú výpočtovo náročné a niekedy nie je nutné estimovať všetky parametre kamery. Nami navrhovaná metóda je vhodná pre estimáciu ohniskových vzdialeností [9, 6]. Budeme vychádzať z modelu perspektívnej projekcie a z predpokladu, že snímaná rovina je umiestnená v rovine hv 3R vizuálnej scény, to znamená, že súradnica  $r$  bude mať nulovú hodnotu. Pre perspektívnu projekciu platí vzťah (2). Budeme uvažovať prípad, kedy sa rotácia  $\mathbf{R}$  a posunutie  $\bar{\mathbf{t}}$  neuplatňujú a teda môžeme písať, že

$$\mathbf{m} = \mathbf{A}\mathbf{M} \quad (5)$$

Čo možno prepísať nasledovne

$$i = -f_y \frac{v}{d-r} + i_0 \quad (6)$$

$$j = f_x \frac{h}{d-r} + j_0 \quad (7)$$

Je zrejmé, že ak sú známe súradnice bodu v priestore a v snímke, potom exaktné riešenie dosiahneme vyjadrením  $f_x$  a  $f_y$  z rovníc (6, 7). Pre ďalšie operácie prepíšeme vzťahy (6, 7) s ohľadom na to, že  $r = 0$  do maticovej formy nasledovne:

$$\begin{bmatrix} -f_y v \\ f_x h \end{bmatrix} - \begin{bmatrix} -d(i - i_0) \\ d(j - j_0) \end{bmatrix} = \bar{\mathbf{0}} \quad (8)$$

V praxi je vhodné riešenie pre viac ako jednu dvojicu korešpondujúcich bodov. Preto ak tieto rovnice prepíšeme do maticového tvaru v zmysle riešenia s minimalizáciou strednej kvadratickej chyby dostávame

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = -(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{B} \quad (9)$$

Kde pre  $n$  korešpondujúcich dvojíc matice  $\mathbf{C}$  a  $\mathbf{B}$  budú mať nasledovný tvar:

$$\mathbf{C} = \begin{bmatrix} X_1 & 0 \\ 0 & Y_1 \\ \vdots & \vdots \\ X_n & 0 \\ 0 & Y_n \end{bmatrix}, \quad \mathbf{B} = d \begin{bmatrix} j_0 - j_1 \\ i_1 - i_0 \\ \vdots \\ j_0 - j_n \\ i_n - i_0 \end{bmatrix}$$

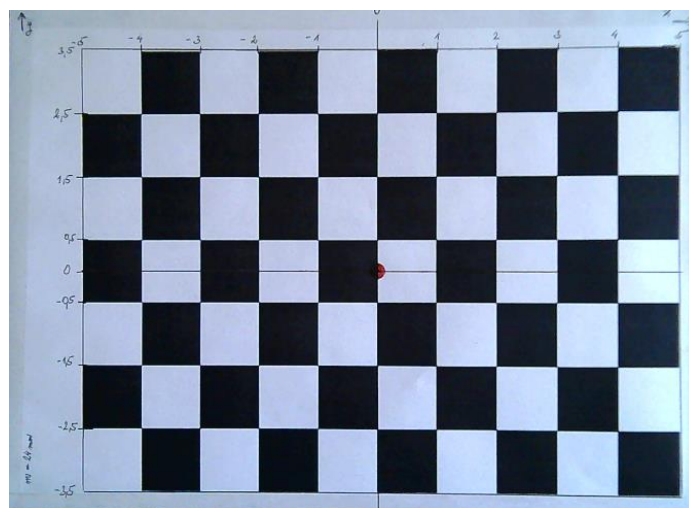
Je zrejmé, že pri takejto estimácii je nutná znalosť vzdialenosti  $d$ , v ktorej je kalibračný obrazec od kamery. Ďalej je zrejmé, že stred snímky nemusí byť uprostred rastra ale bude ním bod, do ktorého sa premietne stred kalibračného obrazca. Tento bod nie je problém odčítať z obrazu. Kvadratickú chybu je možné pre jednotlivé korešpondencie vypočítať nasledovne :

$$\begin{bmatrix} e_y \\ e_x \end{bmatrix} = \left\| \begin{bmatrix} f_y v \\ f_x h \end{bmatrix} - \begin{bmatrix} -d(i - i_0) \\ d(j - j_0) \end{bmatrix} \right\|^2 \quad (10)$$

Pre celkovú kvadratickú chybu bude platiť :

$$e = e_y + e_x \quad (11)$$

Stredná kvadratická chyba  $\bar{e}$  potom bude stredná hodnota všetkých kvadratických chýb. Na obr.3 Je zobrazený kalibračný obrazec nasnímaný zo vzdialenosti 400mm, čo pri mierke 0,25mm /bod predstavuje 1600 bodov.



Obr. 3 Kalibračný obrazec nasnímaný zo vzdialenosti 400mm

V tabl sú uvedené získané ohniskové vzdialenosti pre jednu, 4 a 8 korešpondujúcich dvojíc bodov pre vzdialenosti 400, 600, 800, 1000 mm . Pre kalibráciu budeme využívať body šachovnice so súradnicami :

$$96 \begin{bmatrix} -4 & 2.5 \\ 4 & 2.5 \\ -2 & 1.5 \\ 2 & 1.5 \\ -2 & -1.5 \\ 2 & -1.5 \\ -4 & -2.5 \\ -4 & 2.5 \end{bmatrix}$$

Uvedené súradnice sú prepočítané pomocou bodovej mierky  $m_b = 0.25$  na body. Za stred kalibračného obrazca považujeme bod so súradnicou  $[0 \ 0]$ .

Tabuľka 1. Ohniskové vzdialenosti vypočítané pomocou jednej, štyroch a ôsmich korešpondencií

4 dvojice bodov								
d [mm]	d[op]	$f_x$	$f_y$	$m_{bx}$	$m_{by}$	$\bar{e}$	$l_0$	$J_0$
400	1600	966,74	1024,47	0,25	0,25	11013721953	262,5	449,8
600	2400	944,72	999,87	0,25	0,25	18692282836	281,3	494,1
800	3200	935,28	993,16	0,25	0,25	8489383006	285,4	478,6
1000	4000	934,93	989,8	0,25	0,25	8781077	298,8	449,1
Stredná hodnota:		945,4175	1001,825					
8 dvojíc bodov								
400	1600	964,92	1025,33	0,25	0,25	11131079946	261,875	449,1
600	2400	941,99	1000,66	0,25	0,25	18565565036	282	494,7
800	3200	938,36	992,76	0,25	0,25	8047323623	284,7	477,96
1000	4000	933,18	987,25	0,25	0,25	8334076	299,5	449,8
Stredná hodnota:		944,6125	1001,5					
1 dvojica bodov - exaktný výpočet								
400	1600	957,5	1023,33	0,25	0,25	-	262,5	449,8
600	2400	925,625	1003	0,25	0,25	-	281,3	494,1
800	3200	938,3	992	0,25	0,25	-	285,4	478,6
1000	4000	917,7	996,7	0,25	0,25	-	298,8	449,1
Stredná hodnota:		934,7813	103,758					

Z tab.1 je zrejmé, že ohniskové vzdialenosti sa so vzdialenosťou menia. V skutočnosti takáto zmena nie je možná. Možná by bola iba v prípade, že kamera automatiky zaostruje. V našom prípade však kamera má nastavenú pevnú hĺbkú ostrosť ako aj ohniskovú vzdialenosť. Preto usudzujeme, že zmena ohniskových vzdialeností je spôsobená šumom, ktorý je zavedený kamerou pri snímaní kalibračného obrazu a následne nepresnou voľbou bodov kalibračného obrazca v rovine snímky. Ďalej túto zmenu môže spôsobovať aj skreslenie zavedené optickým systémom kamery. Taktiež chybu je možné zaviesť aj nepresným určením vzdialenosti medzi kamerou a snímaným kalibračným obrazcom. Preto za výslednú hodnotu ohniskovej vzdialenosti budeme považovať strednú hodnotu meraní pri použití 8 korešpondencií.

#### IV. ZÁVER

V príspevku sme popísali matematický postup pri kalibrácii kamery. Nevýhodou klasických techník je to, že pre kalibráciu je nutná aj znalosť hĺbkovej súradnice bodu kalibračného obrazca. Modernjšie metódy bez straty všeobecnosti riešenia využívajú korešpondenciu bodov kalibračnej roviny. Pri tejto metóde nie je nutná znalosť hĺbkovej súradnice bodov korešpondencie. Nevýhodou tejto metódy je však náročnosť výpočtu a nutnosť výpočtu a hľadania homografie medzi snímkami kalibračného obrazca. Nami navrhovaná metóda je zjednodušením prvej a druhej metódy s predpokladom roviny umiestnenej do roviny  $h_v$  s využitím rovníc perspektívnej projekcie. Je však limitovaná tým, že nezohľadňujeme rotáciu a posunutie kamerového systému a taktiež neuvažujeme skreslenia kamerového systému. Nami navrhovaná metóda je vhodná v prípadoch, kedy nie je potrebné vedieť všetky parametre kamerového systému. Touto metódou je možné kalibrovať iba ohniskové vzdialenosti, ktoré sú nutné pri estimácii priestorových súradníc pomocou stereoskopického kamerového systému[10].

#### LITERATÚRA

- [1] ZHANG, Z.: Flexible camera calibration by viewing a plane from unknown orientations, *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on. IEEE, 1999.*
- [2] KIM, J. S., KIM, H. W., KWEON, I. S.: A camera calibration method using concentric circles for vision applications. *ACCV2002, Melbourne, Australia, 2002.*
- [3] ZHANG, Z.: Camera calibration with one-dimensional objects. *Pattern Analysis and Machine Intelligence, IEEE Transactions, 2004.*
- [4] ZHANG, Z.: Camera Calibration, Chapter 2, in G. Medioni and S.B. Kang, eds., *Emerging Topics in Computer Vision, Prentice Hall Professional Technical Reference, 2004.*
- [5] FAUGERAS, O.: *Three-Dimensional Computer Vision: a Geometric Viewpoint, MIT Press, 1993.*
- [6] MIHALÍK, J., MICHALČIN, V.: 3D Motion Estimation and Texturing of Human Head Model. *Radioengineering, Vol.13, No.1, 2004.*
- [7] JIANG, G., ZHAO, C.: *Camera Calibration Based on 2D-plane. 2010..*
- [8] HARTLEY, Richard; ZISSERMAN, Andrew. *Multiple view geometry in computer vision. Cambridge university press, 2003.*
- [9] MIHALÍK, J., MICHALČIN, V.: 3D motion estimation of human head by using optical flow. *Radioengineering, 2006.*
- [10] KOVÁČ, O., MIHALÍK, J.: Estimation of spatial coordinates of 3D objects by stereoscopic scanning, *Acta Electrotechnica et informatica, 2014, (Zaslané redakčnej rade).*

# Finančná predikcia vývoja vybraných menových párov Forex-u

<sup>1</sup> Bc. Róbert Magyar, <sup>2</sup> Ing. Marek Bundzel PhD.

<sup>1</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>2</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>robert.magyar@student.tuke.sk, <sup>2</sup>marek.bundzel@tuke.sk

**Abstrakt** — Predmetom článku je využitie metódy Support Vector Machine na predikciu pohybov menových kurzov a určenie praktického využitia predikcie. Cieľom je výber vhodných menových párov a snaha o nájdenie vhodného modelu a algoritmu, ktorý dokáže z pohybov týchto kurzov profitovať.

**Kľúčové slová** — Forex, SVM, Support Vector Machine, SVM Light, predikcia časových radov, devízový trh, menové páry

## I. ÚVOD

Predikcia menového kurzu na Forexe je zaujímavá téma, ktorá je riešená vďaka jej možným benefitom. Výmena meny za inú menu je dôležitý faktor pri medzinárodnom obchode z hľadiska minimalizácie nákladov a zvyšovania zisku pri stále narastajúcej globalizácii trhov. Zvyšovaním potreby zmeny mien a rozvoja technológií sa vyvinulo niekoľko metód predikcie, ktoré sa líšia hlavne náročnosťou výpočtu, nákladmi, časom výpočtu a presnosťou predikcie. Metódy ako neurónové siete, fuzzy systémy, genetické algoritmy či strojové učenie si našli svoje miesto popri fundamentálnej a technickej analýze, ktoré sú najrozšírenejšie metódy tvoriace základ pre predikciu kurzov na forexovom trhu, trhu s akciami, komoditami a inými.

Problém riešenia predikcie na trhoch je v samotnom správaní sa trhu a jeho účastníkov. Je potreba neustáleho zdokonaľovania a hľadania nových metód, ktoré zvyšujú presnosť predikcie a ktoré môžu pozitívne vplyvať na rozhodnutia vo finančnej oblasti, ktoré sú kľúčovým aspektom v dnešnom štýle biznisu.

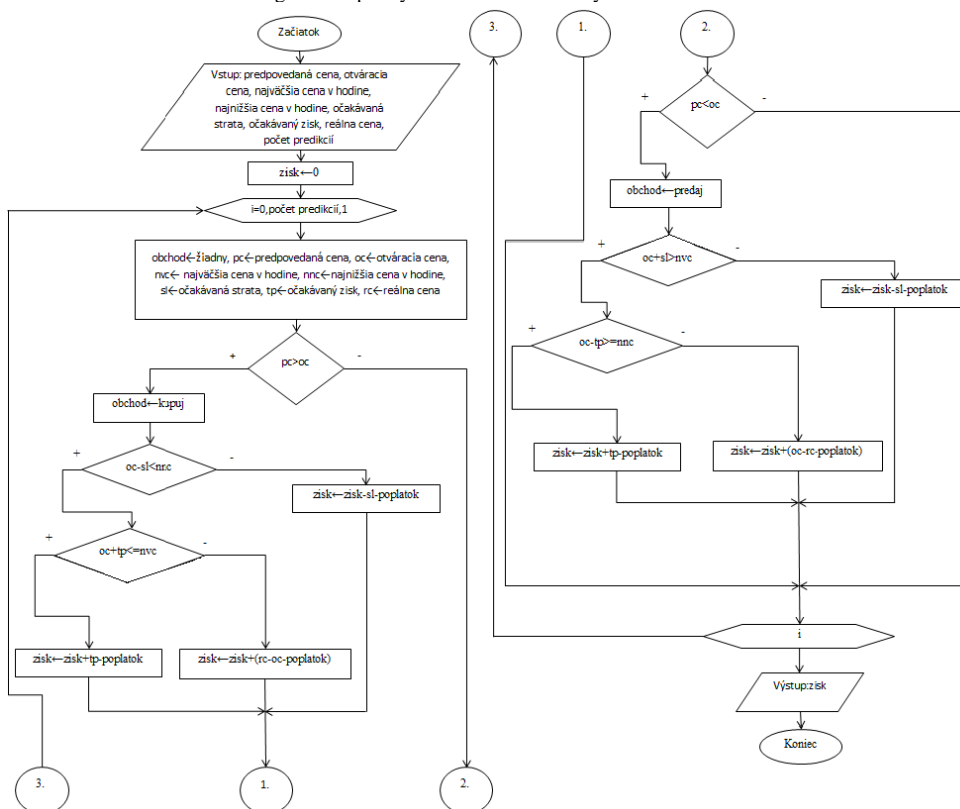
Cieľom experimentov je odpovedať na otázku, či je možné profitovať na forexovom trhu pri využívaní SVM metódy na tvorbu predikcie. Hlavným ukazovateľom úspešnosti je výnosnosť simulovaných obchodov vykonávaná automatickým obchodovaním bez zásahu človeka.

Pre predikciu sme použili metódu Support Vector Machine[4]. Postup riešenia bol v ohodnotení menových párov z hľadiska vhodnosti pre predikciu, kde najdôležitejším kritériom boli náklady na obchod, určenie množstva a typ dát pre predikciu, tréning a otestovanie modelu pri zmene parametrov a funkcií, stanovenie obchodnej stratégie a použitie vybraného modelu na výpočet výnosnosti.

## II. ALGORITMUS PRE VYKONÁVANIE SIMULOVANÝCH OBCHODOV

Pre praktické použitie vybraného modelu je dôležitá simulácia rovnakých podmienok ako pri reálnych obchodoch. Poplatok za obchod sa hradí pri každom nákupe čo zvyšuje riziko obchodu a znižuje potenciálny zisk. Výnos sme vypočítavali na validačných dátach, pričom každá cena predstavovala cenu za jednu hodinu. Pre simuláciu a vykonávanie obchodov sme vytvorili algoritmus, ktorý podľa predikcie kupuje a predáva meny. Algoritmus je popísaný na Obr. 1.

Obr. 1  
Algoritmus pre vykonávanie simulovaných obchodov.



Úlohou algoritmu je rozhodnúť na základe predpovede, či je vhodnejšie predávať alebo kupovať. Po tomto rozhodnutí sa zistí, či sa reálna cena dostala k očakávanému zisku alebo očakávanej strate, ktorú obchodník vopred stanovil. Následne sa vypočíta zisk alebo strata.

### III. EXPERIMENTY

Pre experimenty sme vytvorili aplikáciu, ktorej hlavnou úlohou je automatizovať procesy potrebné pre predikciu. Medzi najdôležitejšie funkcie aplikácie patrí úprava dát, vytvorenie modelu a regresie pomocou funkcií SVM Light[3], nájdenie najlepšieho modelu, reprezentácia dát pomocou grafov, vykonanie iteratívnej predikcie a klasickej predikcie cyklicky, vypočítanie výnosu pomocou simulácie obchodov a vypísanie dodatočných informácií o simulovaných obchodoch.

Experimentálna činnosť bola rozdelená do 3 hlavných fáz. Počiatočná fáza bola v znamení nájdenia a úpravy dát pre SVM metódu, výber funkcie a parametrov pre tvorbu modelu a výber vhodnej metódy pre výpočet chyby predikcie. Hlavná fáza bola v znamení tvorby modelov pri zmene vstupných parametrov a veľkosti vstupných dát. Každému modelu bola určená chyba, teda nepresnosť predikcie v porovnaní s testovacími dátami pomocou vopred zvolenej metódy. Pre model s najmenšou chybou sme vytvorili iteratívnu predikciu a porovnali sme výsledky. V konečnej fáze sme model vyhodnotili, zhodnotili sme výsledky a praktický prínos predikcie.

Vybrané páry pre predikciu boli EUR/USD a USD/JPY. Menové páry sme vybrali z hľadiska minimalizovania poplatkov za obchod. Na tréningovanie modelov sme využívali dáta od novembra 2012 po september 2013. Dáta zo septembra a októbra 2013 sme použili na testovanie modelov a dáta z novembra a decembra 2013 boli použité na validáciu modelu. Využívali sme otváracie, zatváracie, najvyššie a najnižšie ceny vybraných menových párov.

Cieľom bolo nájdenie najvhodnejšieho modelu, kde kritérium vhodnosti bola najmenšia chyba z modelov, s ktorými sa experimentovalo a použitie predikcie vybraného modelu na výpočet výnosu pri simulácii reálnych obchodov. Pred začatím experimentov sme vykonali prípravu dát z hľadiska redukcie a určili sme veľkosť časového okna na vstupe modelu. Pre oba menové páry sa počiatočná množina dát zredukovala o mesiac november z roku 2012, kde sa ukázalo, že tento mesiac neprispieva k presnosti modelu. Tak isto sme zistili, že pre menový pár EUR/USD je optimálna veľkosť časového okna na vstupe modelu 5 menových kurzov, pre pár USD/JPY to boli 2 menové kurzy.

Nájdenie modelu s najväčšou presnosťou predikcie bolo v znamení experimentovania s lineárnou, polynomiálnou a radial basis funkciou a ich parametrami.

Tab. 1  
Porovnanie výkonnosti modelov na dátach menového páru EUR/USD.

Model	MAPE testovacie dáta (EUR/USD)	MAPE validačné dáta (EUR/USD)
RBF: C=40, G=5, W=0,004	0,102	0,1272
LIN: C=20, W=0,001	0,0502	0,0517
POL: C=10, W=0,004, S=30, c=40	0,0538	0,0557

Tab. 2  
Porovnanie výkonnosti modelov na dátach menového páru USD/JPY.

Model	MAPE testovacie dáta (USD/JPY)	MAPE validačné dáta (USD/JPY)
RBF: C=10, G=5, W=0,004	0,7226	19,3525
LIN: C=10, W=0,007	0,0753	0,0668
POL: C=10, W=0,004, S=10, c=50	4,0246	3,9015

Najlepšie dopadla lineárna funkcia pri oboch menových pároch. Najväčší čas zabralo tvorenie predikcie pomocou funkcie RBF pre pár USD/JPY a pomocou polynomiálnej funkcie pre pár EUR/USD. Pre predikciu na EUR/USD je vhodná polynomiálna a lineárna funkcia, radial basis funkcia vzhľadom na presnosť zaostáva. Pre predikciu na páre USD/JPY je použiteľná len lineárna funkcia, ktorá poskytovala zásadne lepšiu presnosť ako ostatné funkcie.

Po nájdení vhodného modelu bolo potrebné optimalizovať parametre očakávaný zisk a riziko z obchodu. Ukázalo sa, že vybraný model priniesol stratu o hodnote -694 pipov na páre EUR/USD a -2142 pipov na páre USD/JPY. Pri neuspokojivých výsledkoch sme zmenili typ predikcie na iterovanú predikciu, pri ktorej sa naraz predikujú dve hodnoty a modifikovali sme algoritmus. Následkom tejto zmeny sa dosiahli o 32 % lepšie výsledky na menovom páre EUR/USD a o 24 % na páre USD/JPY. Keďže výsledky aj po zlepšení boli stratové, bolo potrebné optimalizovať náklady na obchodovanie. Pri vypočítavaní doterajších výnosov sa bral do úvahy poplatok, ktorý si inkasujú slovenské brokerské spoločnosti[1]. Niektoré zahraničné spoločnosti ale poskytujú menšie poplatky (ECN brokeri). Spoločnosť AGM Markets poskytuje na štandardnom účte priemerný poplatok 0,9 pip pre pár EUR/USD a 1,2 pre USD/JPY [2]

Tab. 3  
Vplyv zmeny poplatku na celkový výnos na páre EUR/USD.

Poplatok za obchod v pip	Celkový výnos predikcie (EUR/USD)
1,6	-476
1,3	-181
1,1	15
0,9	211

Tab. 4  
Vplyv zmeny poplatku na celkový výnos na páre USD/JPY.

Poplatok za obchod v pip	Celkový výnos predikcie (USD/JPY)
2,8	-1637
2,0	-855
1,5	-366
1,2	-73

Vysoká frekvencia obchodov znamená kumulovanie nákladov za obchodovanie. Optimalizovanie poplatkov môže zásadne ovplyvniť výnosnosť modelov. Zmenšením poplatku na páre EUR/USD sa strata zmenila na výnos. Ak by sa vykonávali obchody u zahraničnej brokerskej firmy AGM Markets tak by model Lin:C=20 a W=0,001 na páre EUR/USD bol výnosový o hodnote 211 pipov za mesiace november a december 2013. Na páre USD/JPY by sa situácia nezmenila a model by bol aj za týchto podmienok stratový.

#### IV. ZÁVER

Pri experimentoch sa ukázalo, že metóda Support Vector Machine dokáže vytvárať taký model, pomocou ktorého je možné pri určitých podmienkach dosiahnuť na vopred upravených dátach výnosnosť. Pre dosiahnutie výnosnosti je potrebné vybrať najvhodnejšiu funkciu a optimalizovať chybu predikcie pomocou zmeny parametrov. Ukázalo sa, že dôležitým faktorom okrem minimalizácie chyby predikcie je aj spôsob obchodovania, teda zvolený algoritmus a minimalizovanie nákladov.

Kurzy vybraných menových párov EUR/USD a USD/JPY sa správajú zväčša lineárne, keďže regresia pomocou lineárnej funkcie dosiahla najlepšie výsledky. Ukázalo sa, že obchodovanie na menovom páre EUR/USD môže byť profitabilné s navrhnutým algoritmom pri použití iteratívnej predikcie ak náklady na obchod sú pod určitým prahom, bodom kde sa výnos rovná strate. Pre menový pár USD/JPY sa nenašiel broker, ktorý poskytuje také poplatky, ktoré by viedli k výnosu z obchodovania. Na slovenskom trhu sa nenašla brokerská spoločnosť, ktorá by poskytovala také poplatky, s ktorými by predikcia dosiahla výnos. V zahraničí je konkurencia vyššia a pri zvolení určitých ECN brokerov by predikcia na menovom páre EUR/USD priniesla výnos. Pre menový pár USD/JPY sa nenašiel slovenský ani zahraničný broker, ktorý by poskytoval takú veľkosť poplatkov, pri ktorej by bola predikcia zisková.

#### V. POĎAKOVANIE

Táto práca bola vytvorená realizáciou projektu Rozvoj Centra informačných a komunikačných technológií pre znalostné systémy (kód ITMS projektu: 26220120030) na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

#### ZOZNAM POUŽITEJ LITERATÚRY

- [1] T. Joachims, *Advances in Kernel Methods - Support Vector Learning*, MIT-Press, 1999.
- [2] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge, 2000.
- [3] X-Trade Brokers: *Typy obchodných účtov* (online), [cit. 5.11.2013]. Dostupné na internete :<<http://www.xtb.sk/ponuka/basic>>.

# Návrh a implementácia nástroja pre podporu predspracovania dát časových radov

*Dominika LIŠČINSKÁ, Alexandra LUKÁČOVÁ, Ján PARALIC*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická Univerzita v Košiciach, Slovensko

dominika.liscinska@student.tuke.sk, alexandra.lukacova@tuke.sk,  
jan.paralic@tuke.sk

**Abstrakt** — Tento článok sa zaoberá opisom návrhu a implementácie interaktívnej webovej aplikácie na podporu predspracovania dát časových radov. Vzhľadom na široké využitie časových radov v rôznych vedných oblastiach vzniká potreba aplikácie na vhodné predspracovanie veľkého množstva zaznamenaných časových radov pre analytické a výskumné účely. Cieľom navrhutej a implementovanej aplikácie je nielen umožniť používateľom bez znalosti programovacieho jazyka predspracovať dáta pomocou základných funkcií ako vyhľadanie časového radu, spriemerovanie či vyhľadávanie lokálnych extrémov, ale najmä poskytnúť vhodnú transformáciu časového radu na symbolické atribúty pre analýzu napr. prostredníctvom hľadania sekvenčných vzorov formou asocičných pravidiel. Návrh aplikácie zohľadňuje požiadavky jednoduchosti použitia jednotlivých funkcií a zabezpečenia vhodného výstupu novozískaných dát.

**Kľúčové slová** — aplikácia, časový rad, predspracovanie, programovací jazyk R, Shiny

## I. ÚVOD

Predspracovanie dát je nevyhnutnou činnosťou pri časových radoch s veľkým množstvom zaznamenaných údajov. Navrhnutá a vytvorená aplikácia slúži jednak používateľom bez programátorských znalostí na rýchle a efektívne predspracovanie časových radov a jednak výskumníkom, ktorí si potrebujú predspracovať dáta pre sekvenčnú analýzu nepravidelných časových dejov. Medzi jej hlavné charakteristiky patrí jednoduchosť použitia, interaktivita a voľná dostupnosť.

Na základe získaných teoretických poznatkov o časových radoch boli navrhnuté štyri základné funkcie: vizualizácia, vyhľadanie časového radu kľzavým priemerom, výpočet aritmetického priemeru a vyhľadávanie dôležitých bodov s následnou transformáciou symbolickú reprezentáciu, ktorá umožňuje ďalšiu analýzu ako napr. hľadanie sekvenčných vzorov bez presného vzorkovania, ktoré v mnohých reálnych aplikáciách nie je účelné. Aplikácia je naprogramovaná v programovacom jazyku R, ktorého prídavný balík Shiny, umožňuje vytvárať moderné webové aplikácie.

Nasadenie aplikácie v slovenskom a anglickom jazyku na voľne dostupnom webovom portáli umožňuje prístup všetkým potenciálnym používateľom. Zverejnenie zdrojového kódu na internetovej stránke poskytuje používateľom príležitosť pre prípadné rozšírenie, alebo modifikáciu aplikácie.

Zvyšok tohto príspevku je rozdelený nasledovne: v kapitole 2 prezentujeme analýzu súčasného stavu. V kapitole 3 rozoberáme návrh aplikácie. Detailným popisom funkcií vyvinutej aplikácie sa zaoberáme v kapitole 4 a nasadeniu aplikácie venujeme kapitolu 5.

## II. ANALÝZA SÚČASNÉHO STAVU

### A. Časový rad

Podľa [1] môžeme časový rad charakterizovať ako postupnosť hodnôt jednej sledovanej veličiny zaznamenávanej v čase. Časové rady, zvyčajne zobrazované pomocou grafov, sú využívané v mnohých odboroch, v ekonomike sú to napríklad denné zmeny kurzov na burze, v medicínskej oblasti to môže byť magnetická rezonancia zaznamenávajúca činnosť



mozgových vln. Anderson [2] uvádza štyri zložky časového radu: trendová, sezónna, cyklická a reziduálna (náhodná) zložka.

### B. Analýza časových radov

Analýza časových radov nám pomáha pochopiť vývoj sledovanej veličiny a predpovedať jej budúce hodnoty. Pochopenie fungovania časového radu umožňuje vytvoriť taký model, ktorý dokáže monitorovať, predpovedať a kontrolovať daný jav [3]. Pri zaznamenávaní hodnôt časového radu je získavané veľké množstvo dát, ktoré je potrebné pred ich ďalším skúmaním spracovať. Práve na tento účel slúži navrhnutá aplikácia.

### C. Súčasné možnosti

Existuje množstvo nástrojov súvisiacich s problematikou časových radov. Many Eyes [4] je webová aplikácia, ktorá slúži na vykreslenie a zdieľanie grafickej reprezentácie užívateľových dát. Výhodou tejto aplikácie je široká možnosť výberu zobrazenia grafu. Yeroon.net/ggplot2 [5] je webová aplikácia pre vykreslenie načítaných dát zo súborov. Na jej použitie nie je potrebné ovládať programovanie. Táto aplikácia bola vytvorená pomocou programovacieho jazyka R. Jej ovládanie je pomerne jednoduché a intuitívne. Wessa.net [6] obsahuje kalkulátory na analýzu časových radov. Dáta sa dajú jednoducho načítať a vybraný kalkulátor vypočíta výsledok a vykreslí danú situáciu do grafu. Táto aplikácia taktiež vytvorená pomocou programovacieho jazyka R.

Napriek tomu, že je dostupných mnoho aplikácií, väčšina sa zameriava iba na vykreslenie časových radov do grafov. Za nevýhodu preto považujeme absenciu možnosti ich ďalšieho spracovania. Samotný programovací jazyk R ponúka veľa možností na prácu s časovými radmi. Problémom pre mnohých užívateľov je však nutnosť ovládania programovacieho jazyka. Taktiež je používanie programovacieho jazyka pre používateľov veľmi zdĺhavé a daný čas by mohli využiť efektívnejšie.

## III. NÁVRH APLIKÁCIE

### A. Programovací jazyk R

Aplikácia je navrhnutá v programovacom jazyku R, ktorý je známy ako štatistický jazyk. Medzi jeho hlavné výhody patrí jednoduchosť, flexibilita a množstvo prídavných balíkov rozširujúcich funkcionality jazyka.

### B. Shiny

Shiny [7] je jedným z prídavných balíkov programovacieho jazyka R, ktorý slúži na vytváranie interaktívnych webových aplikácií, a preto bol vhodným nástrojom pre účel vytvorenia navrhovanej aplikácie. Pri používaní balíka Shiny nie je potrebné ovládať iné programovacie jazyky ako R. Moderný grafický dizajn aplikácií balíka Shiny je zaručený využívaním projektu Twitter Bootstrap.

### C. Funkcionalita aplikácie

Aplikácia obsahuje štyri základné funkcie: vizualizácia dát, funkcia kľzavého priemeru, aritmetického priemeru a funkcia vyhľadávania dôležitých bodov s následnou možnosťou transformácie časového radu na množinu symbolických atribútov viazaných na nájdené významné body. Funkciu aritmetického priemeru je možné prepojiť s kľzavým priemerom a teda spriemerovať vopred vyhladené dáta. Okrem týchto základných funkcií ponúka aplikácia stiahnutie novovzniknutých dát vo forme súboru s dátami ako aj grafu. Pod každým vykresleným grafom jednotlivých funkcií sa nachádza aj sumár základných štatistických údajov z novozískaných dát.

## IV. FUNKCIE APLIKÁCIE

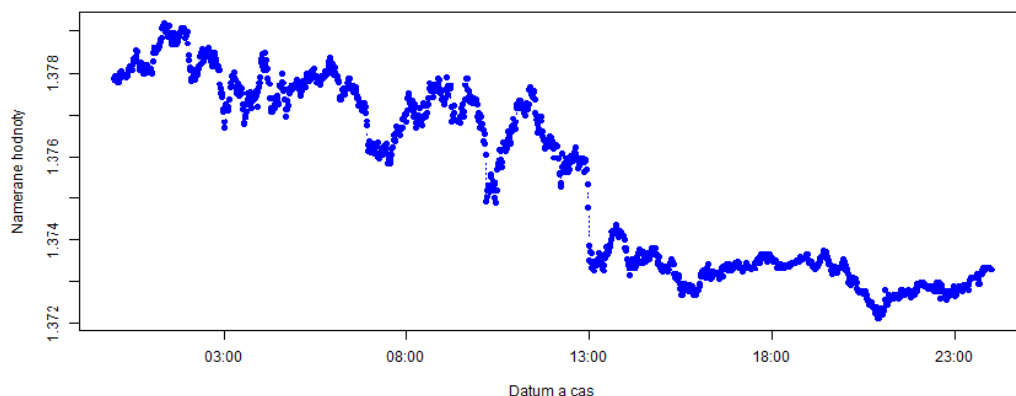
### A. Vizualizácia dát

Funkcia vizualizácie dát je najjednoduchšou funkciou aplikácie. Slúži na vykreslenie načítaných dát zvoleného časového radu v grafe (viď Obr.1).

### A. Kľzavý priemer

Metóda kľzavého priemeru slúži na očistenie časového radu od jeho sezónnych a náhodných zložiek. Aplikácia vyrovnáva časový rad na základe používateľom zadanej dĺžky kľzavej časti. Čím je dĺžka kľzavej časti väčšia, tým viac je časový rad vyrovnanjší. Používateľ musí nájsť pre daný účel použitia vhodnú dĺžku kľzavej časti pre vybraný časový rad tak, aby bol rad dostatočne vyhladený, no zároveň neprišiel o dôležité body zlomu vývoja časového radu. Práve

v tom mu je nápomocná navrhnutá aplikácia, nakoľko môže okamžite sledovať efekt každej zmeny spomínaného parametra.



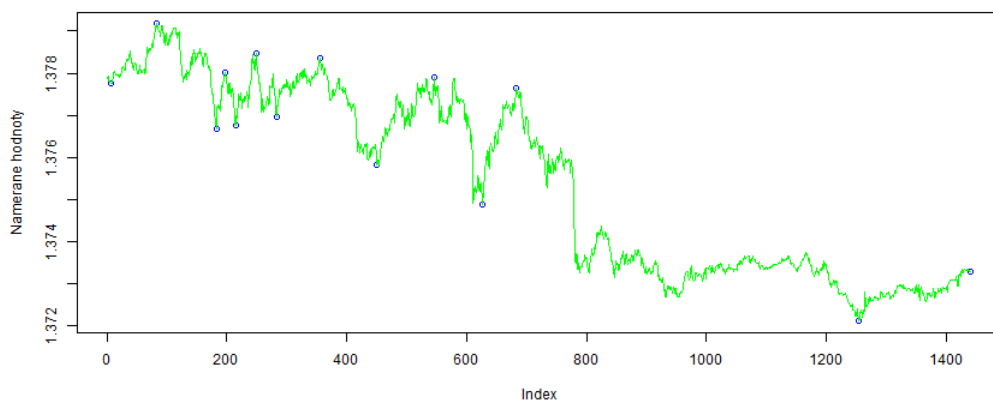
Obr.1: Ukážka vizualizácie dát

### B. Aritmetický priemer

Funkcia aritmetického priemeru slúži na spriemerovanie údajov zaznamenaných v krátkych časových intervaloch pre získanie priemernej hodnoty za zvolený časový úsek. Používateľ zadáva frekvenciu priemeru, pričom má možnosť spriemerovať buď pôvodný, alebo už vopred vyhladený časový rad.

### C. Hľadanie dôležitých bodov

Funkcia hľadania dôležitých bodov je určená na vyhľadávanie lokálnych extrémov časového radu. Algoritmus funkcie bol naprogramovaný podľa pseudokódu uvedeného v knihe Finka a Pratta [8]. Používateľ zadáva veľkosť kompresného pomeru R, ktorý hovorí o tom, aké veľký rozdiel medzi jednotlivými hodnotami nás zaujímajú. Čím menší kompresný pomer R, tým viac nájdených dôležitých bodov a naopak, pričom R musí byť vždy väčšie ako 1 (vid' Obr.2).



Obr.2: Ukážka funkcie vyhľadávania dôležitých bodov

Pod grafom s vykreslenými nájdenými dôležitými bodmi sa nachádza tabuľka so symbolickou transformáciou nájdených hodnôt. V tabuľke je uvedený dátum, čas a hodnota nájdeného dôležitého bodu, informácia o tom či bolo dosiahnuté lokálne maximum alebo minimum, aký stav pretrvával do daného bodu a ako dlho tento stav trval (vid' Obr.3).

	Datum	y	Extrem	Stav	Trvanie_v_min
1	2014-03-03 00:06:00	1.37775	lok. min	-	-
2	2014-03-03 01:22:00	1.37917	lok. max	stupa	76
3	2014-03-03 03:02:00	1.37668	lok. min	klesa	100

Obr.3: Ukážka symbolickej transformácie nájdených lokálnych extrémov

## V. NASADENIE APLIKÁCIE

Aplikácia je nasadená v anglickom aj slovenskom jazyku pomocou riešenia ShinyApps<sup>1</sup>, ktoré umožňuje pre nekomerčné využitie cloud hosting Shiny aplikácií zdarma. Anglická verzia je navyše obohatená o množinu dát, keď si chce užívateľ aplikáciu iba vyskúšať. Nachádza sa na stránke <https://projekty.shinyapps.io/Application/>. Zdrojový kód v anglickom jazyku je zverejnený na stránke <https://github.com/DominikaLiscinska/App>.

## VI. ZÁVER

Výsledné riešenie spĺňa vopred dané požiadavky a obsahuje základné funkcie na predspracovanie dát časových radov. Pomocou prídavného balíka Shiny programovacieho jazyka R sa podarilo vytvoriť jednoduchú interaktívnu webovú aplikáciu s moderným grafickým dizajnom. Naprogramovanie aplikácie v jazyku R, ktorý je známy využívaním najmä v štatistike, umožní používateľom jednoduché rozšírenie funkcií aplikácie, prípadne ich modifikáciu. Nasadenie aplikácie v slovenskom a anglickom jazyku na voľne dostupnom portáli umožní ľahký prístup všetkým používateľom. Zverejnenie zdrojového kódu umožňuje modifikáciu a rozšírenie funkcionality aplikácie pre účely konkrétneho výskumu.

## POĎAKOVANIE

Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

## ZOZNAM POUŽITEJ LITERATÚRY

- [1] G. E. P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2013.
- [2] T.W. Anderson, *The Statistical Analysis of Time Series*. John Wiley & Sons, 2011.
- [3] D. Liščinská, *Návrh a implementácia aplikácie pre podporu predspracovania dát časových radov*. 2014.
- [4] IBM Advanced visualization. *Many Eyes*. 2013.
- [5] OOMS, Jeroen C.L. A web interface for the R package ggplot2. Verzia 0.2., 2010.
- [6] Wessa, P. Free Statistics Software. Verzia 1.1.23-r7, Office for Research Development and Education, 2013.
- [7] RStudio, Inc. *Web Application Framework for R*. 2014.
- [8] E. Fink, K. B. Pratt, *Indexing of Compressed Time Series*. Computer Science Department, 2004.

---

<sup>1</sup> <http://www.rstudio.com/products/hosted-shiny/>

# Analýza možností výberu dát z databázy ORACLE

<sup>1</sup>Veronika VARGOVÁ, <sup>2</sup>Matej ČOPÍK, <sup>3</sup>Ján JADLOVSKÝ

<sup>1,2,3</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>veronika.vargova@tuke.sk, <sup>2</sup>matej.copik@tuke.sk, <sup>3</sup>jan.jadlovsky@tuke.sk

**Abstrakt** — Tento článok sa zaoberá analýzou možností výberu dát z databázy ORACLE a vytvorením používateľského prostredia na zadávanie SQL príkazov, ich vyhodnotenie a zobrazenie. Zadaný SQL príkaz je spracovaný pomocou databázových procedúr, ktorých výsledkom je zoznam alternatívnych príkazov. Na vyhodnotenie týchto príkazov sa použil nástroj EXPLAIN PLAN. Príkazy boli spustené na dvoch typoch testovacieho prostredia. Najprv nad tabuľkou s nagenеровanými údajmi v počte troch miliónov. Potom nad databázou pre Flexibilný montážny podnik. Aplikácia zobrazuje zoznam výsledkov spolu s informáciou o nákladoch a čase spracovania.

**Kľúčové slová** — databázový systém, optimalizácia, výber dát

## I. ÚVOD

Článok poukazuje na možnosti spracovania SQL príkazov. Vytvoriť taký SQL príkaz, ktorý by predstavoval vysokú efektivitu pri spracovávaní je častokrát veľmi komplikované. Vysokú zložitosť tomu pripisuje množstvo variácií, ktoré je možné v rámci napísania jedného príkazu vytvoriť. Cieľom je zjednodušiť a zrýchliť možnosti, ako pretransformovať jeden výraz na taký, ktorý bude pre dané použitie vhodnejší. Používateľ prostredníctvom vytvorenej aplikácie získa rôzne formy spracovania pre ním zadaný príkaz. Ďalšou pridanou hodnotou aplikácie je, že ponúkne nielen možné variácie zadaného príkazu, ale aj dodatočné informácie.

Tieto informácie sa budú venovať času spracovania daného príkazu, a zároveň to budú údaje o nákladoch, ktoré vynaloží počítač pri spracovávaní príkazu. Vytvorená aplikácia umožňuje vytváranie nových príkazov na základe komunikácie s databázou Oracle. Na základe tejto komunikácie pripravuje konkrétne transformácie pre zadaný výraz. Tie budú umožňovať používateľovi usporiadanie výsledných príkazov podľa vlastnej potreby, a zároveň pre lepšiu interpretáciu získaných výsledkov bude možné vygenerovať graf.

## II. OPTIMALIZÁCIA SQL PRÍKAZOV

Pod pojmom „Query optimalizácia“ sa rozumie schopnosť napísať SQL príkaz tak, aby bol vykonaný čo najrýchlejšie, pri čo najmenšom zaťažení systému. Optimalizácia sa vníma ako jedna z funkcií relačných databázových systémov. Jednotlivé príkazy SQL ponúkajú kombinácie, ako ich implementovať, aby sa dosiahlo ich optimálne využitie. Pri programovaní sa používajú základné pravidlá, ktorých dodržiavanie prináša rýchlejšie vykonávanie jednotlivých príkazov. Tieto pravidlá sú definované primárne pre Oracle. Jednotlivé pravidlá, ako zlepšiť SQL príkaz sú uvedené v nasledujúcich častiach.

### A. Indexy v tabuľkách

Počet indexov použitých v tabuľkách musí byť správne určený, pretože veľký počet indexov môže spôsobiť menšiu efektivnosť práce. Indexy ukladané do databázy sú definované ako nezávislé objekty. Index je podobný tabuľke – obsahuje riadky. Neodporúča sa vytvárať indexové súbory na tabuľkách, ktoré majú menej ako 1500 riadkov. Ak zrušíme tabuľku, tak sa automaticky zrušia aj všetky jej indexové súbory. Vždy po vložení nových záznamov do tabuľky sa automaticky aktualizujú aj jej indexy.

Odporúča sa indexovať atribúty, ktoré sa často používajú spolu s kľúčovým slovom SQL jazyka WHERE. Atribúty, ktoré sa používajú na spojenie medzi tabuľkami je rovnako vhodné indexovať. Unikátne indexy sa používajú pre atribúty primárneho kľúča a neunikátne pre cudzí

klúč, ktoré sú často používané spolu s WHERE. Indexy konštruujeme takou metódou, aby indexované atribúty nasledujúce po WHERE boli prvé. Atribúty, ktoré sa často menia sa neindexujú. Príkazy INSERT, UPDATE A DELETE spomaľujú výkon servera, pretože potrebujú aktualizovať indexy.

#### B. Funkcia UPPER

Funkcia UPPER patrí medzi reťazové funkcie. Je to funkcia, ktorá nám vracia reťazec. Jej význam je definovaný tak, že zmení všetky malé písmena na veľké. Nedovoľuje používanie skenovania po indexoch, a preto prebehne celkové prezeranie tabuľky.

```
SELECT DEPT_NAME FROM DEPARTMENT  
WHERE UPPER(DEPT_NAME) like 'SALES%'
```

#### C. Výber správnej varianty pri výbere dát z tabuliek

Je dôležité sa rozhodnúť, či sa bude skenovať cez index alebo tabuľky sa budú prezerat globálne. Je potrebné si uvedomiť, že platí, ak je výber dát väčší ako 20% všetkých dát z tabuľky, tak je globálne prezeranie rýchlejšie ako skenovanie cez index.

Globálne prezeranie:

```
SELECT * +FULL  
FROM EMP  
WHERE SALARY = '1000'
```

Inštrukcia FULL núti ORACLE server k prezeraniu celej tabuľky. Skenovanie cez index sa nepoužíva ani v prípade, ak ORACLE server musí vykonávať sekundárnu transformáciu údajov, čo znamená, že symbolická hodnota sa transformuje do číselnej.

#### D. Dosiahnutie efektívnosti

Efektívnosť sa môže dosiahnuť rôznymi metódami. Existujú operátory pri ktorých platí, že pracujú rýchlejšie ako iné operátory. Patrí tu napríklad operátor MINUS, ktorý vykoná príkaz rýchlejšie, ako operátory s WHERE NOT IN (SELECT) alebo WHERE NOT EXISTS.

```
SELECT CUST_ID FROM CUSTOMERS  
WHERE STATE IN ('SK','CZ')  
AND CODE NOT IN (804,410)
```

Predchádzajúce volanie sa prepíše s operátom MINUS, ako je na nasledujúcom volaní, a následne je možné používať aj indexové skenovanie.

```
SELECT CUST_ID FROM CUSTOMERS  
WHERE STATE IN ('SK','CZ')  
MINUS  
SELECT CUST_ID FROM CUSTOMERS  
WHERE CODE IN (804,410)
```

#### E. Použitie špeciálnych stĺpcov

Riadok ROWID predstavuje najrýchlejší prístup k dátam. Ako príklad je uvedený operátor UPDATE, ktorý používa operátor skenovania ROWID.

```
SELECT ROWID, SALARY INTO TEMP_ROWID, TEMP_SALARY  
FROM EMPLOYEE
```

### III. APLIKÁCIA

Cieľom vytvorenej aplikácie je analyzovať možnosti výberu dát z databázy Oracle. Samotná aplikácia je navrhnutá tak, že komunikuje pomocou obsluhy v programovacom jazyku C#. Pri návrhu aplikácie boli teoreticky spracované možnosti výberu dát z databázového systému. Aplikácia je navrhnutá tak, že vykonáva činnosť, až po pripojení na databázu. V stave, ak je používateľ pripojený k databáze je možné s aplikáciou vyhľadávať možné optimalizačné prístupy, pri spracovávaní zadaného príkazu.

To znamená, že používateľ do zadávacieho okna vloží príkaz, a následne spustí proces spracovania príkazu. Program spustí proces zlepšovania a vráti používateľovi získané

výsledky. Výsledky sú používateľovi zobrazené vzhľadom na čas trvania, a vzhľadom na celkové náklady spracovania. Tieto výsledky vracia používateľovi nástroj, ktorý nám ponúka Oracle 11g, a to EXPLAIN PLAN. Používateľ si môže zvoliť vlastné kritéria zoradenia výsledkov. Aplikácia mu ponúka možnosť zvolenia si vhodného usporiadania. A to, buď podľa rýchlosti, alebo podľa najnižšieho zaťaženia. Pre lepšie grafické zobrazenie výsledkov ponúka aplikácia možnosť vygenerovania grafu. Používateľ si rovnako môže vybrať, čo chce mať vo vytvorenom grafe zobrazené, resp. môže porovnávať obe kritéria.

#### A. Testovacia databáza

Zdrojom údajov potrebných na analýzu sú vygenerované údaje pomocou použitia vlastného generátora. Vygenerovali sa údaje v objeme troch miliónov záznamov. Údaje sú vo forme najpoužívanejších dátových typov.

	↕ COLUMN_NAME	↕ DATA_TYPE	↕ NULLABLE	DATA_DEFAULT	↕ COLUMN_ID	↕ COMMENTS
1	ID	NUMBER	Yes	(null)	1	(null)
2	ZNAKY_PEVNA	CHAR(1000 BYTE)	Yes	(null)	2	(null)
3	ZNAKY_VAR	VARCHAR2(3000 BYTE)	Yes	(null)	3	(null)
4	NZNAKY_PEVNA	NCHAR(1000 CHAR)	Yes	(null)	4	(null)
5	NZNAKY_VAR	NVARCHAR2(2000 CHAR)	Yes	(null)	5	(null)
6	CISLO	NUMBER(38,10)	Yes	(null)	6	(null)
7	DATUM	DATE	Yes	(null)	7	(null)
8	DATUM_CAS	TIMESTAMP(6)	Yes	(null)	8	(null)

Obr. 1: Zdrojová tabuľka pre testovacie údaje

V prvej fáze testovania bol pomocou aplikácie spracovávaný jednoduchý typ príkazu. Pod jednoduchým príkazom sa rozumie príkaz, ktorý obsahuje v klauzule WHERE len jednoduché viacnásobné podmienky. V podmienkach sa pracuje s operátormi: NOT LIKE, BETWEEN – AND, znak rovnosti (=) a znak nerovnosti (< alebo >). SQL príkaz bol zadaný v nasledujúcom tvare:

```
SELECT MAX(cislo) FROM datetypes
WHERE znaky_pegna NOT LIKE 'traot'
AND cislo BETWEEN 34 AND 5897
AND nznaky_pegna = '%LKA'
AND znaky_var != 'PI_KK'
GROUP BY TRUNC(datum)
```

Po zadaní príkazu a stlačení *OPTIMIZE* aplikácia zobrazí možné variácie pôvodného príkazu. Rovnako pri každej variácii je možné nájsť aj informácie o *COST* a *TIME*, ktoré sú vyhodnotené pomocou *EXPLAIN PLAN*. Výsledky sú usporiadané podľa nákladov zostupne.

Po vložení príkazu sa vygenerovali možné alternatívy pôvodného príkazu. V ponúkaných možnostiach bolo vytvorených 64 možných príkazov, a to vrátane pôvodného príkazu zadaného používateľom. Medzi jednotlivými výsledkami boli zaznamenané výrazné rozdiely pri hodnotách nákladov, menšie pri hodnotách času.

P. č. príkazu	SQL príkaz	Priemerný čas	Interval
5	select max(cislo) from datetypes where znaky_pegna <> 'traot' and cislo between 34 and 5897 and nznaky_pegna = '%lka' and znaky_var != 'pi_kk' group by trunc(datum)	123,680	(121,262;125,998)
6	select max(cislo) from datetypes where znaky_pegna <> 'traot' and cislo between 34 and 5897 and nznaky_pegna = '%lka' and znaky_var ^= 'pi_kk' group by trunc(datum)	123,704	(122,439;124,968)
63	select max(cislo) from datetypes where znaky_pegna <> 'traot' and cislo between 34 and 5897 and nznaky_pegna = '%lka' and znaky_var <> 'pi_kk' group by trunc(datum)	135,860	(122,623;140,568)

Okrem sledovania štatistických údajov je potrebné brať do úvahy aj účel daného SQL príkazu. Napriek tomu, že sa na prvých miestach pohybuje príkaz s poradovým číslom 5 a 6, nie je možné ich odporúčať ako najvhodnejšie, a to z jedného dôvodu. Vo „WHERE“ klauzule sa nachádza podmienka so stĺpcom, ktorý je znakového typu, a obsahuje tzv. žolíky („%“ a „\_“). Tieto žolíky vie spracovať len príkaz „LIKE“, iné príkazy považujú žolíky za obyčajné znaky. Preto výsledky získané s použitím príkazu „LIKE“ sú iné ako s inými operátormi.

### B. Databáza Flexibilného montážneho podniku (FMP)

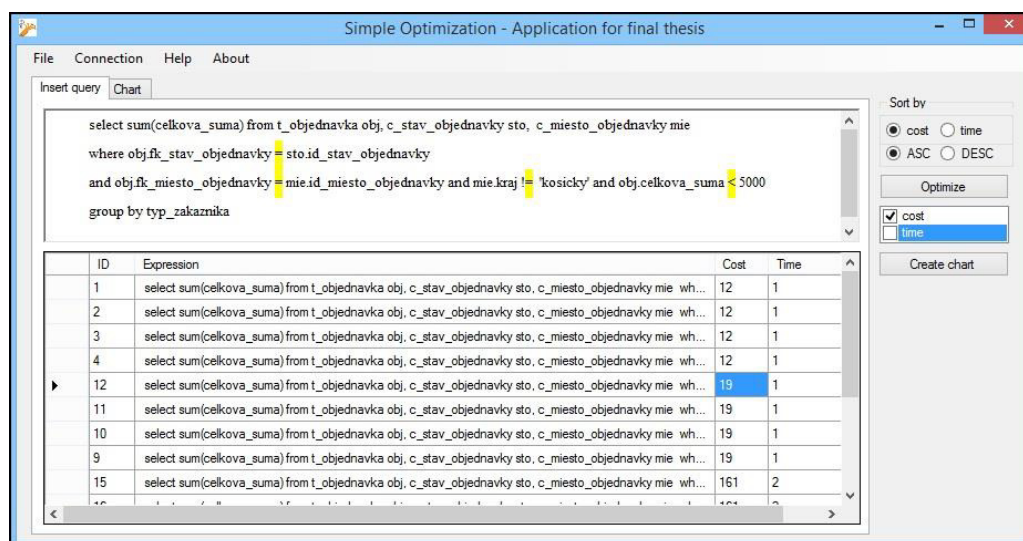
Zdrojom údajov pre druhú časť optimalizácie SQL príkazov je databáza pre FMP. Príkazy vyberajú dáta z tabuliek, ktoré obsahujú informácie o objednávkach a produktoch výrobnéj linke. Aj v tejto časti testovania bolo cieľom nájsť nové alternatívne podoby príkazov pre používateľom zadaný príkaz. V rámci týchto alternatív sa získavali informácie, ktoré nám poskytuje nástroj EXPLAIN PLAN.

Prvý príkaz má za úlohu vybrať z databázy súčet hodnôt objednávok pre všetky typy zákazníkov, ktorí nepochádzajú z košického kraja s hodnotou objednávky menšou ako 5000.

Zadaný príkaz bol definovaný v nasledujúcom tvare:

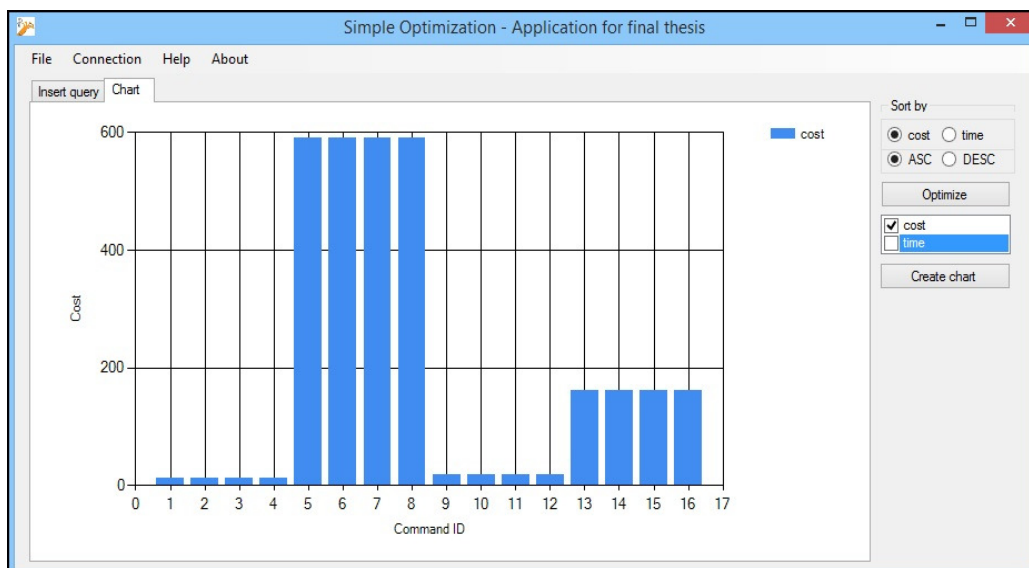
```
SELECT SUM(celkova_suma) FROM t_objednavka obj, c_stav_objednavky sto,
        c_miesto_objednavky mie
        WHERE obj.fk_stav_objednavky = sto.id_stav_objednavky
        AND obj.fk_miesto_objednavky = mie.id_miesto_objednavky
        AND mie.kraj != 'kosicky'
        AND obj.celkova_suma < 5000
        GROUP BY typ_zakaznika
```

Pomocou aplikácie bol spracovaný SQL príkaz a z pôvodného príkazu sa vytvorilo 16 rôznych alternatív. Jednotlivé príkazy boli zotriedené podľa nákladov zostupne.



Obr. 2: Alternatívne SQL príkazy vygenerované aplikáciou

Pôvodný SQL príkaz je možné považovať za optimálny, aplikácia navrhla použiť vo „WHERE“ klauzule aj iné rovnocenné operátory nerovnosti, avšak dosiahli sa rovnaké výsledky ako pri pôvodnom SQL príkaze. Zaujímavé hodnoty sú pri tých alternatívach, pri ktorých sa aplikácia pokúšala nahradiť operátor rovnosti „=“ operátorom „LIKE“. Operátor rovnosti je v príkaze použitý pri spájaní tabuliek medzi stĺpcami, ktoré sú primárnymi kľúčmi tabuliek a navyše aj dátového typu „NUMBER“. Takýto spôsob zápisu SQL príkazu sa považuje za veľmi nevhodný, čo potvrdzuje aj graf na obr. 3, kde hodnota nákladov dramaticky vyskočila na niekoľkonásobné hodnoty, čo viedlo aj k zvýšeniu hodnoty pre čas spracovania.



Obr. 3 Zobrazenie nákladov jednotlivých príkazov

#### IV. ZÁVER

Nielen čas spracovania je pri jednotlivých vytvorených príkazoch dôležitý, ale rovnako aj náklady vzťahujúce sa na spracovanie príkazu. Náklady sa vzťahujú na využitie CPU, pamäte, ako aj spotrebu elektrickej energie. Pretože, energetika a otázka šetrenia energií je otázkou budúcnosti, sú informácie o nákladoch dôležitým smerníkom pri výbere príkazu. Samozrejme je stále v rukách používateľa, akú cestu považuje pre výkon svojej činnosti za prvoradú. Môže to byť cesta najrýchlejšieho spracovania príkazov, alebo v inom prípade je to cesta ušetrenia nákladov pri zadávaní príkazu. Rovnako aj v tomto prípade existuje tzv. „zlatá stredná cesta“, kedy je možné vybrať pomalšie spracovanie príkazu a lepšie hodnoty vynaložených nákladov. Úlohou aplikácie je uľahčiť toto rozhodovanie a priniesť základný prehľad o alternatívach k používanému SQL príkazu.

#### V. POĎAKOVANIE

„Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.“

#### REFERENCIE

- [1] DATE, C.J.: An Introduction to Database Systems, USA: PEARSON, 1990. 983s. ISBN 032-119-784-4
- [2] DATE, C.J.: Twelve Rules for a Distributed Database, 1987. s. 77-81.
- [3] FEADSON: Introduction to Database System, [online]. 2010, [cit. 2014-03-10]. Dostupné na internete: <<http://books.google.sk/books?id=y7P9sa2MeGIC&printsec=frontcover&dq=database+systems&hl=sk&sa=X&ei=91cwU4rGJIbR4QTAs4DABg&ved=0CGYQ6AEwBw#v=onepage&q=database%20systems&f=false>>.
- [4] ZANILOLO, Carlo: Advanced Database System. USA: San Francisco, 1997. 575 s. ISBN 1-55860-443-X.
- [5] GARMANY, J. – HARTMANN, L. – CARR, B. – JAIN, V.J. – KARAM, S.: Oracle 11g: New Features: Get Started Fast with Oracle 11g Enhancements. USA: Kittrell, 2008. 468 s. ISBN 978-0-9797951-0-7.



# Návrh a realizácia riadenia modelu inkubátora pre podporu rastu húb pomocou jednočipového mikropočítača

<sup>1</sup>Matúš KONKOL, <sup>2</sup>Matej ČOPÍK, <sup>3</sup>Ján JADLOVSKÝ

<sup>1</sup> Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>2,3</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>matus.konkol@student.tuke.sk, <sup>2</sup>matej.copik@tuke.sk, <sup>3</sup>jan.jadlovsky@tuke.sk

**Abstrakt** — Tento článok popisuje skonštruovanie a uvedenie do prevádzky modelu inkubátora, ktorý je určený primárne na stimuláciu rastu vlákniťých húb. Model upravuje jeho vnútorné prostredie dostupnými periférnymi zariadeniami tak, aby vytvoril v čo najväčšej miere ideálne podmienky pre rast organizmov v ňom umiestnených. Konfigurácia je realizovaná užívateľom pomocou dotykovej klávesnice a aktuálne hodnoty sú zobrazené na znakovom displeji. Umožňuje pripojenie k počítaču a vytvorenie záznamu jeho prevádzky, kedy je možná kontrola jeho správania.

**KLúčové slová** — Inkubátor, vlákniťá, huba, riadenie, stimulácia, rast, mikropočítač

## I. ÚVOD

Počas svojej existencie si človek vďaka svojej inteligencii osvojil získavanie potravy rôznymi technikami. Tie vytvoril na základe sledovania prirodzeného prostredia rastlín a živočíchov, ktoré sa snažil napodobňovať. V dnešnej dobe sú možnosti v oblasti automatizácie procesov vďaka použitiu mikropočítačov takmer neobmedzené. V úvodnej časti sú popísané procesy návrhu a výroby modelu. Ďalej sa nachádza popis programu riadiacej jednotky inkubátora a programu na záznam údajov do počítača a nakoniec popis nameraných hodnôt z reálnej prevádzky.

## II. MIKROBIOLOGICKÝ INKUBÁTOR

Inkubátor alebo kultivačný box je zariadenie, ktoré kontroluje a riadi isté dôležité parametre pre vývoj rôznych organizmov v uzavretom priestore oddeleného od exteriéru. Snahou je docieľenie čo najideálnejších podmienok pre rast a vývoj daného organizmu.

Huby ako mikrobiologický organizmus, pre ktorý je tento model určený, sú zložité a náročné na podmienky, v ktorých dokážu rásť a množiť sa. Práve rozmnožovanie je sprevádzané vývojom plodníc, ktoré by mali byť výsledkom inkubácie húb. Počas nepriaznivého obdobia sa uvádzajú do hibernácie, v ktorej dokážu vytrvať veľmi dlhú dobu. [1] Účel inkubátora je hibernácii predchádzať a to kontrolovaním podmienok, v ktorých sa tieto organizmy nachádzajú. Taktiež ich podporovať vo vývoji a rozmnožovaní, čo vedie k typickému prejavu húb počas vegetačného obdobia – rastu plodníc, ktoré musia byť bez kontaminantov, aby ich bolo možné využívať na prípadnú konzumáciu.

## III. NÁVRH A VÝROBA MODELU

### A. Telo inkubátora

Model inkubátora pozostáva z pohyblivých a nepohyblivých pasívnych častí a elektroniky, ktorá je riadená mikropočítačom. Ako termoizolačný materiál, z ktorého pozostáva telo inkubátora, je použitý polystyrén. Ten umožňuje dostatočnú elimináciu ovplyvňovania vnútorného prostredia tým vonkajším. Povrchová úprava v podobe cementového práškoveho lepidla spevňuje jeho povrch, ktorý je tak spevnený a inak krehký polystyrén je chránený voči

mechanickému poškodeniu.

Telo ma tvar kvádru, kde sa v prednej časti nachádza otvor tvaru obdĺžnika, ktorý je určený na manipuláciu s objektmi nachádzajúcimi sa vo vnútornom priestore inkubátora (Obr. 1). Otvor sa uzatvára dvierkami, ktoré je možné zafixovať, aby sa zlepšili izolačné vlastnosti modelu. Vo vnútornom priestore sa nachádzajú deliace prepážky, ktoré fyzicky členia vnútorný priestor a predstavujú podklad pre nádoby s mikrobiologickým obsahom.

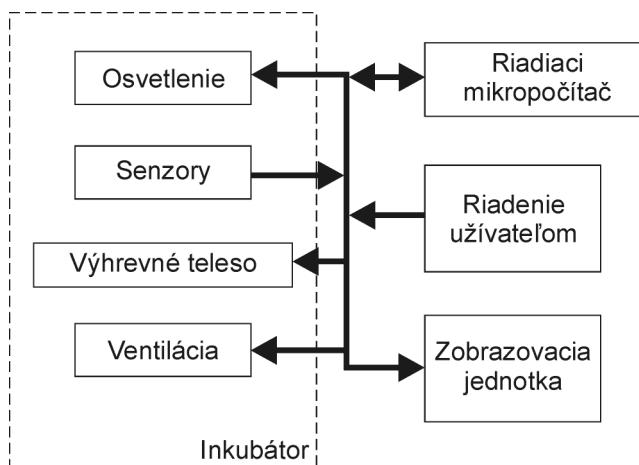


Obr. 1 Inkubátor ošetrový povrchovou úpravou s uzatváracími dvierkami

### B. Periférie

Periférne zariadenia (Obr. 2), ktoré slúžia na ovplyvňovanie vnútorného priestoru pozostávajú so senzora DHT22, ktorý sníma teplotu a vlhkosť prostredia, v ktorom sa nachádza. Ten zabezpečuje presné merania, ktorého hodnoty slúžia ako vstupy pre algoritmy riadenia.

Riadenie teploty je zabezpečené odporovým drôtom, ktorý umožňuje vyhrievanie vnútorného prostredia po privedení napätia. Stabilné zvlhčovanie prostredia je realizované pomocou nádob naplnených perlitom s vodou. Tento spôsob predstavuje statické a neustále odparovanie vody vďaka jej zväčšenému povrchu, ak je to potrebné. LED diódy slúžia na osvetlenie priestoru v zhoršených svetelných podmienkach a stimuluje mycélium huby k produkcii plodníc. [1] Ventilátory na zadnej strane zariadenia odstraňujú nahromadený oxid uhličitý vytvorený organizmami a vŕhá vzdušný kyslík cez filtre slúžiacie ako ochrana pred kontamináciou. Signalizačné LED diódy umiestnené na vrchnej strane riadiaceho obvodu indikujú funkčnosť jednotlivých periférií.



Obr. 2 Bloková schéma zapojenia

### C. Riadiaci obvod

Jadrom riadiaceho obvodu je mikropočítač PIC18F2550 [4], ktorý zabezpečuje komunikáciu jednotlivých komponentov a logiku celého procesu riadenia. Presnosť meraného času zabezpečujú obvody DS1305 (RTC – hodiny reálneho času [5]) a DS32KHz (kryštál s teplotnou

kompenzáciou). Čas predstavuje jeden zo vstupov pre algoritmy riadenia, preto je neustále aktualizovaný. Jednotlivé elektrotechnické súčiastky sú umiestnené na dvojstrannej doske plošných spojov, ktorá je chránená plastovým puzdrom. Pod povrchom inkubátora sú vedené elektrické vodiče, ktoré zabezpečujú prepojenie všetkých komponentov a ku riadiacemu obvodu sú pripojené pomocou kolíkovej lišty.

#### D. Interakcia s používateľom

Keďže je inkubátor zariadením, ktoré má poskytovať ideálne podmienky pre rôzne vývojové štádia organizmu huby, musí byť vybavený ovládacími prvkami. Tie pozostávajú so štandardného 14 znakového dvojriadkového znakového displeja a dotykovej klávesnice, ktoré sú umiestnené v ľahko dostupnej časti zariadenia vedľa riadiaceho obvodu. Prepojenie s ním je realizované pomocou plochého kábla, ktorý sa napája na dosku plošných spojov umiestnenú pod displejom. Klávesnicu predstavuje predloha 14-tich tlačidiel na papieri umiestnená pod rezistívnou dotykovou plochou. Polohu dotyku je vďaka použitej technológii možné vyčítať ako hodnotu napätia na jednotlivých elektródach pre obe osy. Displej obsahuje podsvietenie, ktoré je aktivované iba v prípade, kedy sa s inkubátorom pracuje. Obsahuje štandardnú znakovú sadu, ktorá je rozšírená o znaky s diakritikou.

### IV. POPIS PROGRAMOV

#### A. Program mikropočítača

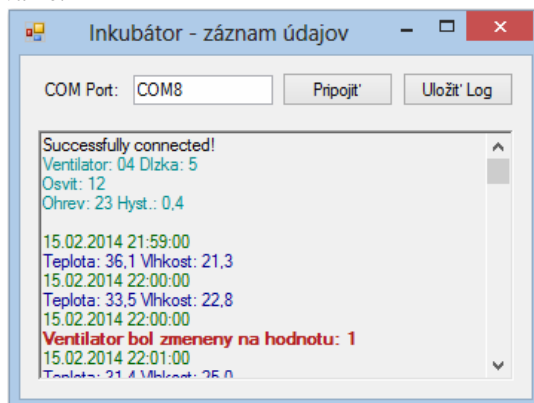
Program je písaný v jazyku assembler vo vývojovom prostredí MPLab, ktoré vyvíja firma Microchip, ktorej mikropočítač je použitý v modeli. Po skompilovaní bol nahraný do mikropočítača pomocou nástroja PicKit 2. Vďaka vyvedeným kontaktom určeným na programovanie je možné program kedykoľvek aktualizovať napríklad pomocou hore spomenutých nástrojov.

Tvorba aplikácie pozostávala z prototypu, kedy bola funkčnosť algoritmov a jednotlivých periférií kontrolovaná na kontaktnom poli, čo v značnej miere uľahčilo vývoj. Vďaka použitému programovaciemu jazyku nebolo možné efektívne program rozdeliť do niekoľkých súborov a preto sa nachádza v jedinom, kde sú jednotlivé logické bloky oddelené komentármi pre jednoduchšiu orientáciu. Komunikácie boli realizované softvérovo pomocou pomenovaných vývodov z mikropočítača, čo vo veľkej miere uľahčilo následný návrh dosky plošných spojov, nakoľko bolo možné funkcie týchto vývodov jednoducho vymeniť.

Hlavný program pozostáva z algoritmov riadenia jednotlivých periférií, ktoré zabezpečujú ich správnu funkčnosť počas celého behu. Ďalšie funkcie, ako čítanie času, prekresľovanie obsahu displeja, získavanie údajov so senzora a podobne sú spravované časovačmi, ktoré vyvolávajú prerušenia.

#### B. Program na záznam údajov

Program je písaný v jazyku C# na ktorého vývoj bolo použité vývojové prostredie Visual Studio od Microsoftu. Umožňuje čítanie záznamu prevádzky inkubátora, ktorý sa pripojí k počítaču pomocou sériového kábla (Obr. 3). Zaznamenanie času, dátumu, teploty a vlhkosti sa realizuje každú minútu a v prípade, že sa zmenilo napájanie periférií. V takejto situácii sa vypíše aj text oznamujúci túto skutočnosť. Údaje je možné následne uložiť ako obyčajný text, alebo ako údaje oddelené čiarkou. Následne môžu byť tieto informácie uložené do databázy, prípadne použité na ďalšie spracovanie.

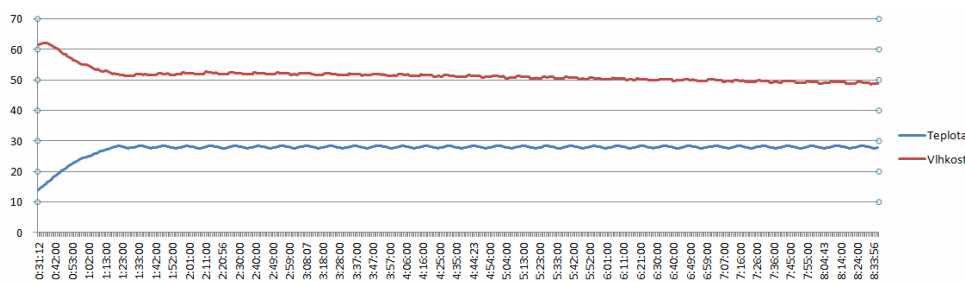


Obr. 3 Program na záznam údajov

## V. EXPERIMENTÁLNE MERANIA

Po skonštruovaní zariadenia bolo možné overiť jeho funkčnosť. Testy boli realizované na tomto zariadení v rôznych konfiguráciách a podmienkach, ktoré by mohli ovplyvniť jeho chod. Overenie funkčnosti je dôležité z hľadiska použiteľnosti inkubátora na účel, pre ktorý bol vyrobený. Merania sa týkali sledovania vlhkosti, teploty a správania sa podľa definovanej konfigurácie, kedy boli testované všetky periférie. Po overení správnej funkcionality bol inkubátor podrobený testovaniu s mycéliom huby, ktorej rast bol stimulovaný vplyvom podmienok, ktoré inkubátor vytvára.

Overenie správnej funkčnosti inkubátora spočívalo v jeho dlhodobom spustení a monitorovaní vlastností vnútorného prostredia pomocou programu na záznam údajov. Tieto údaje boli uložené vo formáte s dátami oddelenými čiarkou a bol z nich vytvorený graf, ktorý v závislosti od času zobrazuje priebeh teploty a vlhkosti vnútorného prostredia (Obr. 4). Po uskutočnení niekoľkých meraní bol vyvodенý záver, že sa zariadenie správa podľa očakávaní a je možné ho použiť na inkubovanie húb.



Obr. 4 Priebeh teploty a vlhkosti v závislosti od času v priestore inkubátora

## VI. ZÁVER

Rast a rozmnožovanie mikroorganizmov je zložitým biologickým procesom. Rozmnožovanie húb je spojené s tvorbou plodníc, ktoré sú po tisíročia používané ako potrava nielen ľuďmi. Správna stimulácia týchto organizmov ich podnecuje k rýchlejšiemu rastu a priblíženie sa ideálnym podmienkam môže spôsobiť napríklad vytvorenie väčšieho množstva plodníc. [2] Kontrolu týchto podmienok realizuje inkubátor, ktorého vnútorný priestor je monitorovaný mikropočítačom a prostredie je upravované periférnymi zariadeniami. Počas dlhodobších meraní bola preukázaná správna funkčnosť modelu, a teda je možné použiť inkubátor na experimenty s hubovým organizmom.

## VII. POĎAKOVANIE

„Tento príspevok vznikol vďaka podpore v rámci operačného programu Výskum a vývoj, pre projekt: Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií, kód ITMS: 26220220182, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.“

## LITERATÚRA

- [1] HUDECOVÁ, Daniela a Viktor MAJTÁN. Mikrobiológia I. 1. vyd. Bratislava: STU, 2002, s. 47-78. ISBN 80-227-1663-4.
- [2] JABLONSKÝ, Ivan a Václav ŠAŠEK. Jedlé a léčivé houby: pěstování a využití. Vyd. 1. Praha: Brázda, 2006, 263 s., [16] s. barev. obr. příl. ISBN 80-209-0341-0.
- [3] Nahuby.sk. BARANOVIČ, Roland. Pestovanie hlívy ustricovej [online]. [cit. 2013-11-03]. Dostupné z: [http://www.nahuby.sk/clanok.php?clanok\\_id=15](http://www.nahuby.sk/clanok.php?clanok_id=15).
- [4] MICROCHIP, Technology. PIC18F2455/2550/4455/4550 Data Sheet [online]. 2006 [cit. 2013-11-03]. Dostupné z: <http://ww1.microchip.com/downloads/en/devicedoc/39632c.pdf>.
- [5] Maxim Integrated. DS1305 Serial Alarm Real-Time Clock. [online]. [cit. 2014-04-05]. Dostupné z: <http://datasheets.maximintegrated.com/en/ds/DS1305.pdf>

# Ohrev DPS vnorenými rezistormi

*Lubomír LIVOVSÝ*

Katedra technológií v elektronike, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

lubomir.livovsky@tuke.sk

**Abstrakt** — Článok sa zaoberá možnosťou ohrevu dosky plošných spojov (DPS) vnorenými rezistorov na báze NiP (nikel fosforovej) odporovej vrstvy. Navrhnuté riešenie ohrievača je možné využiť pre potreby stáleho ohrevu DPS s možnosťou zabezpečenia konštantnej pracovnej teploty pre kritické elektronické prvky a pod. Maximálna teplota, ktorú je možné pomocou vnorených rezistorov dosiahnuť je daná vlastnosťami materiálov, z ktorých je DPS vyrobená.

**Kľúčové slová** —vnorený rezistor, ohrev, DPS

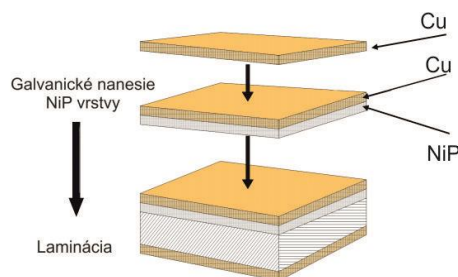
## I. ÚVOD

Pasívne súčiastky predstavujú veľkú časť elektronických zariadení. Nové zariadenia s väčšou integráciou funkcií, menšou veľkosťou a nižšími výrobnými nákladmi vyžadujú menšie, kompaktnéjšie a zložitejšie konštrukčné riešenia. Jednou z možností ako zmenšovať rozmery zariadenia je umiestnenie pasívnych súčiastok do vnútra substrátu napr. dosky plošného spoja. Takýmto riešením je možné zmenšovať rozmery zariadenia, znížiť náklady na výrobu, znížiť vplyv parazitných účinkov na kvalitu signálov. Vnorené rezistory je možné použiť ako ukončovacie a prispôsobovacie rezistory, „pull up“ rezistory a podobne. Neštandardné využitie vnorených rezistorov ponúka plošné alebo lokálne generovanie tepla. Využitie vnoreného rezistora, ktorý je súčasťou DPS je vhodné pre zabezpečenia konštantnej pracovnej teploty na ploche DPS, čím je možné eliminovať vplyv zmeny okolitej teploty na presnosť merania analógových veličín.

## II. MATERIÁL PRE VNORENÉ REZISTORY

Vytvorenie vnorených rezistorov je podmienené použitím odporového materiálu, ktorý je možné laminovať do výslednej zostavy DPS. Materiál musí vyhovovať technológii výroby viacvrstvovej DPS, čím sa proces výroby DPS rozšíri. Jednou z možností ako vytvoriť vnorený rezistor je použitie materiálu na báze nikel-fosforovej (NiP) zliatiny, ktorá vykazuje vysokú stabilitu hodnôt odporu a vyhovuje procesu technológie výroby DPS a taktiež procesu spájkovania súčiastok bezolovnatou spájkou.

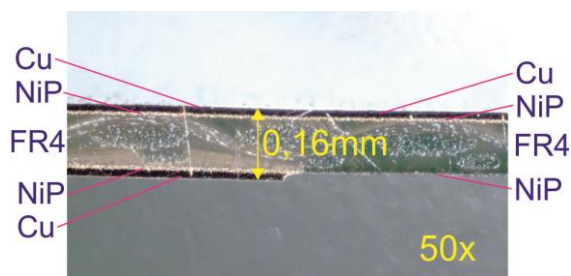
Materiál na báze nikel-fosforovej zliatiny s označením „OhmegaPly“ ponúka riešenie pre tvorbu vnorených rezistorov vo vnútorných vrstvách viacvrstvovej DPS. Nikel-fosforová zliatina a medená fólia vytvoria tenkú odporovo-vodivú vrstvu označovanú ako RCM (Resistor-Conductor Material). RCM materiál je laminovaný s dielektrickým materiálom (epoxidovým, polyimidovým, PTFEs), čím je možné vytvoriť dosku plošného spoja s vnútornou odporovou vrstvou (Obr.1). Týmto spôsobom je možné vytvoriť základnú dvojstrannú DPS na báze FR4 s dvomi vodivými vrstvami a jednou alebo dvomi NiP odporovými vrstvami [1].



Obr. 1: OhmegaPly - RCM materiál.

Z materiálu OhmegaPly sa procesom viacnásobného leptania vytvárajú vnorené rezistory.

Konečná viacvrstvomá DPS môže podľa potreby obsahovať niekoľko vrstiev základného materiálu. Podľa hrúbky NiP vrstvy je možné vytvoriť RCM materiál s rôznymi hodnotami plošného odporu ( $R_s$  - odpor na jednotku plochy).



Obr. 2: Rez DPS po vyleptaní vnoreného rezistora

Na obrázku Obr. 2 je zobrazený rez FR4 dosky plošného spoja s vyleptaným vnoreným rezistorom. Pod oboma vrstvami medi sa nachádzajú vrstvy NiP, ktoré spolu so sklenenými vlákňami a epoxidovou živickou tvoria RCM materiál. Na spodnej strane DPS je vidieť odleptanú med', čím bol vytvorený vnorený rezistor.

Tabuľka 1  
 Oblasti použitia vnorených rezistorov

Plošný odpor $R_s$	Tolerancia	Hrúbka NiP	Aplikácia, použitie
10 $\Omega/\square$	3%	1.00 $\mu\text{m}$	Sériové ukončovacie rezistory, ktoré sú súčasťou vodivej čiary.
25 $\Omega/\square$	5%	0.40 $\mu\text{m}$	Používané pre sériové a paralelné ukončenie a výkonové odporové deliče.
50 $\Omega/\square$	5%	0.20 $\mu\text{m}$	Použitie ako „Pull-up“, „Pull-down“ rezistory pre logické obvody.
100 $\Omega/\square$	5%	0.10 $\mu\text{m}$	
250 $\Omega/\square$	10%	0.05 $\mu\text{m}$	Rezistory s vysokou hodnotou odporu.

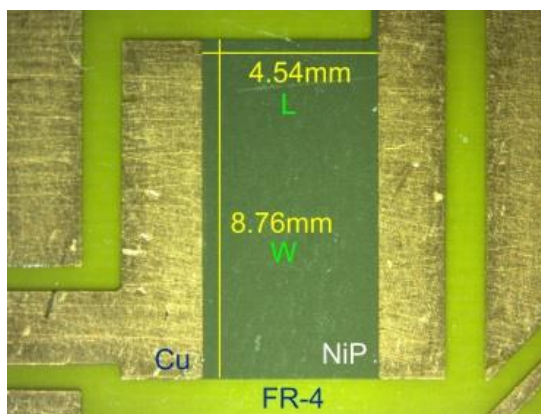
Hodnoty plošného odporu RCM a oblast' použitia vnorených rezistorov sú uvedené v tabuľke Tabuľka 1.

### III. NÁVRH ODPORU VNORENÉHO REZISTORA

Plošný odpor ( $R_s$ ) je odpor na jednotku plochy a nezávisí na veľkosti strany štvorca rezistora. Pri návrhu rozmerov a tvaru požadovaného rezistora sa vychádza z hodnoty plošného odporu použitej NiP vrstvy. Hodnota odporu rezistora v tvare obdĺžnika sa vypočíta podľa vzťahu

$$R = R_s \frac{L}{W} [\Omega], \quad (1)$$

kde  $R$  je požadovaná hodnota rezistora,  $R_s$  – odpor na jednotku plochu,  $L$  – dĺžka rezistora,  $W$  – šírka rezistora (Obr. 3).



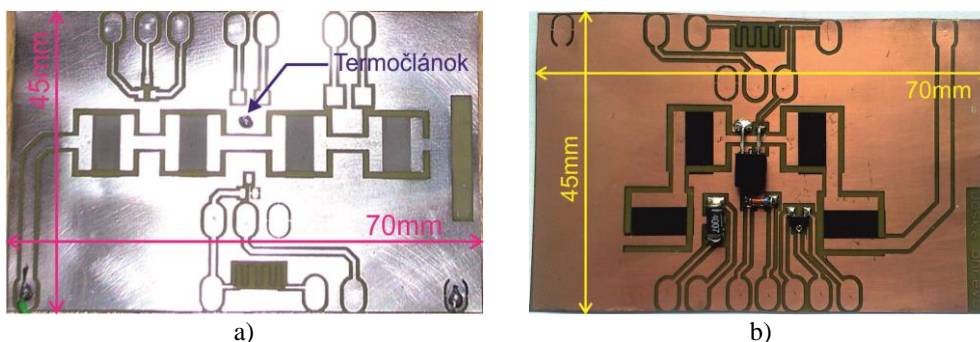
Obr. 3: Rozmery vyhrievacieho rezistora

Veľkosť plochy výkonového rezistora je určená hlavne hodnotou stratového výkonu na



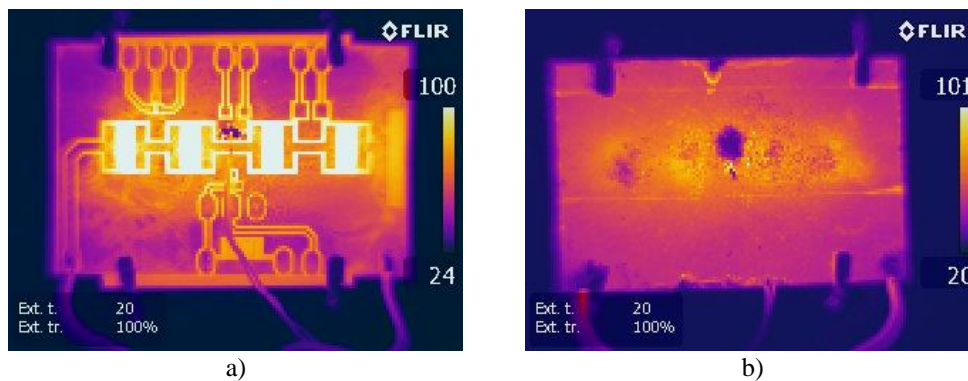
rezistore. Pretože vnorený rezistor je súčasťou substrátu (laminátu), fyzikálne a teplotné charakteristiky substrátov majú dôležitý vplyv na určenie veľkosti plochy výkonového rezistora. Stratový výkon vnoreného rezistora je ovplyvnený: plochou rezistora, hrúbkou a typom DPS, usporiadaním viacvrstvovej štruktúry DPS, okolitou teplotou, tepelnou vodivosťou DPS, prídavnými chladiacimi prvkami napr. veľkosť plochy medi v okolí rezistora.

Pre potreby ohrevu DPS boli zvolené rezistory s hodnotou odporu  $5 \Omega$  a použitý bol RCM materiál s plošným odporom  $R_s = 10 \Omega/\square$  a hrúbkou  $0,16 \text{ mm}$ . Rozmery rezistorov boli vypočítané zo vzťahu (1) s ohľadom na maximálny stratový výkon  $5 \text{ W}$ . Z dôvodu maximálnej pracovnej teploty ( $140^\circ\text{C}$ ) materiálu FR4 nebol pri meraniach prekročený stratový výkon  $2,25 \text{ W}$ . Usporiadanie výhrevných rezistorov je vidieť na obrázku Obr. 4. Úlohou týchto rezistorov je ohrievať DPS s rozmermi  $70 \text{ mm} \times 45 \text{ mm}$  maximálne na teplotu  $120^\circ\text{C}$  za účelom udržiavanie konštantnej teploty (termostat) na ploche DPS. Vyrobené boli dva typy DPS s rôznym rozmiestnením vyhrievacích rezistorov (Obr. 4).



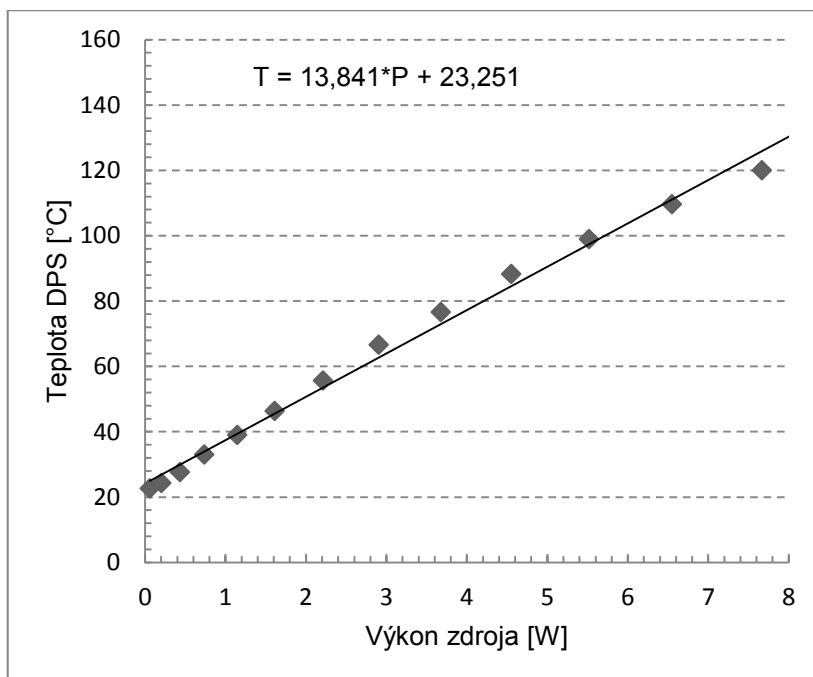
Obr. 4: DPS s vnorenými rezistormi a) povrchová úprava - cín, b) povrchová úprava - lak

Testy ohrevu pomocou rezistorov boli vykonané na DPS FR4 - RCM s hrúbkou  $0,16 \text{ mm}$  a na DPS, ktorá bola vytvorená spojením FR4 – RCM a jednostrannej DPS FR4 s hrúbkou  $1,5 \text{ mm}$ , pričom vyhrievacie rezistory boli vo vnútri vytvorenej štruktúry. Na obrázku Obr. 5 je vidieť rozloženie teploty v oboch prípadoch pričom termočlánok, ktorým sa merala teplota, bol umiestnený v strede DPS.



Obr. 5: Rozloženie teploty na DPS a) pohľad na rezistory hrúbka DPS  $0,16 \text{ mm}$  b) pohľad na Cu vrstvu hrúbka výslednej DPS  $1,66 \text{ mm}$

Závislosť teploty DPS (v mieste termočlánku) a výkonu napájacieho zdroja je zobrazená na obrázku (Obr. 6). Z priebehu grafu vidieť lineárnu závislosť. Maximálna teplota DPS pri teplote okolia  $22^\circ\text{C}$  bola  $120^\circ\text{C}$  s výkonom zdroja  $7,7 \text{ W}$  (prúde  $0,6 \text{ A}$ , napätie  $13 \text{ V}$ ).



Obr. 6: Závislosť teploty DPS na výkone napájacieho zdroja

#### IV. ZÁVER

Testy ohrevu DPS pomocou vnorených rezistorov ukázali, že je možné dosiahnuť teplotu na DPS, ktorá by nedegradovala samotnú DPS. Vhodným rozložením rezistorov je možné získať rovnomerné rozloženie teploty. Spojením vyhrievacích odporov s regulátorom teploty je možné získať termostat pre meranie volt - ampérovej charakteristiky polovodičových súčiastok v procese výučby predmetov v bakalárskom stupni štúdia, čím by sa nahradil ohrev pomocou transformátorového oleja. Zmena teploty DPS je lineárna a testy ukázali, že táto závislosť sa nemení ani pri opakovanom viacnásobnom ohrievaní.

#### POĎAKOVANIE



Podporujeme výskumné aktivity na Slovensku. Projekt je spolufinancovaný zo zdrojov EÚ. Tento článok bol vypracovaný v rámci projektu "Centrum excelentnosti integrovaného výskumu a využitia progresívnych materiálov a technológií v oblasti automobilovej elektroniky", ITMS 26220120055.

#### LITERATÚRA

- [1] OhmegaPly: Detailed Processing Guide [online]. 2008 [cit 2013-05-29]. Dostupné na internete: [http://www.ohmega.com/%3Fs2member\\_file\\_download\\_F-5-Overview-Processing-Guide.pdf](http://www.ohmega.com/%3Fs2member_file_download_F-5-Overview-Processing-Guide.pdf)
- [2] R. N. Das, K. I. Papatomas, J. M. Lauffer, S. Rosser, M. D. Poliks and V. R. Markovich: Manufacturing Substrates with Embedded Passives, The PCB Magazin, November 2011, str. 11-15.
- [3] LIVOVSÝ, Ľubomír: Vnorené rezistory v Altium Designeri. In: DPS - Elektronika od A po Z. roč. 3, č. 3 (2013), s. 28-31.
- [4] OhmegaPly: Design Guide [online]. 2010 [cit 2014-03-25]. Dostupné na internete: [http://www.ohmega.com/member\\_file\\_download=F-4-Overview-Design-Guide.pdf](http://www.ohmega.com/member_file_download=F-4-Overview-Design-Guide.pdf)
- [5] Weiyang Qiu, Kailin Pan, Dongmei li, Jing Liu: Study on Process of Embedded Passives Based on FR4 Substrate Series, 2010 11th International Conference on Electronic Packaging Technology & High Density Packaging, str. 815-818.