

Technical University of Košice



**Faculty of Electrical Engineering  
and Informatics**

**Electrical Engineering and Informatics**

**XII**

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

**ISBN 978-80-553-3940-5**

2021

**Electrical Engineering and Informatics XII**  
**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, 040 01 Košice, Slovak Republic

Date of publication: August 2021 Language: English, Slovak  
Printing: 50 pieces CD Pages: 471

Editorial board chairman: prof. Ing. Alena Pietriková, CSc.

Proceedings reviewers: prof. Ing. Ján Paralič PhD.  
prof. Ing. Daniela Perduková, PhD.  
prof. Ing. Alena Pietriková, CSc.  
prof. Ing. Ján Šaliga, CSc.  
Dr.h.c. prof. RNDr. Ing. Ján Turán, DrSc.  
prof. Ing. Iveta Zolotová CSc.  
assoc. prof. Ing. Norbert Ádám, PhD.  
assoc. prof. Ing. Anton Baláž, PhD.  
assoc. prof. Ing. Lubomír Beňa, PhD.  
assoc. prof. Ing. Marek Bundzel PhD.  
assoc. prof. Ing. Peter Butka PhD.  
assoc. prof. Ing. Lubomír Doboš, PhD.  
assoc. prof. Dr. Ing. Bystrík Dolník  
assoc. prof. Ing. Miloš Drutarovský, PhD.  
assoc. prof. Ing. Jaroslav Džmura, PhD.  
assoc. prof. Ing. Mária Gamcová, PhD.  
assoc. prof. Ing. Ján Gamec, PhD.  
assoc. prof. Ing. Eva Chovancová, PhD.  
assoc. prof. Ing. Ján Kaňuch, PhD.  
assoc. prof. Ing. Iraida Kolcunová, PhD.  
assoc. prof. Ing. Milan Lacko, PhD.  
assoc. prof. Ing. Marián Mach CSc.  
assoc. prof. Ing. Kristína Machová PhD  
assoc. prof. Ing. Dušan Medveď, PhD.  
assoc. prof. Ing. Ján Papaj, PhD.  
assoc. prof. Ing. Jaroslav Petráš, PhD.  
assoc. prof. Ing. William Steingartner, PhD.  
assoc. prof. Dr. Ing. Ján Vaščák  
assoc. prof. Ing. Jaroslava Žilková, PhD.

Editors: prof. Ing. Alena Pietriková, CSc.  
Ing. Emília Pietriková, PhD.

# Contents

<b>Martin Kolek, Roman Cimbala, Peter Havran</b> <i>Electrophysical properties of solid insulation materials</i> .....	7
<b>Luboš Ovseník, Jakub Oravec, Maroš Lapčák, Norbert Zdravecký</b> <i>Návrh automatizovaného systému pre tvrdé prepínanie hybridnej FSO/RF linky na báze strojového učenia</i> .....	12
<b>Luboš Ovseník, Jakub Oravec, Norbert Zdravecký, Maroš Lapčák</b> <i>Návrh 32-kanálového Ultra-DWDM systému pre potlačenie vplyvu nelineárnych javov SPM, XPM a FWM</i> .....	18
<b>Daniel Florek, Anton Baláž, Branislav Madoš</b> <i>Rozšírenie Minecraftu z pohľadu fungovania Bitcoinu</i> .....	24
<b>Branislav Sobota, Sára Javorková, Marián Hudák, Štefan Korečko</b> <i>The therapy of upper limbs with virtual reality technology</i> .....	38
<b>Andrea Petrášová, Branislav Madoš, Anton Baláž</b> <i>Návrh grafického súborového formátu pre ukladanie dvoj-dimenzionálnych dát s využitím HDS</i> ..	45
<b>Tara Stefányi, Valerie Novitzká</b> <i>Categorical model of simply-typed lambda-calculus with reference type</i> .....	51
<b>Matúš Križánek, František Babič, Viera Pavlišková</b> <i>Vytvorenie prototypu na spracovanie transakčných dát v MSP pomocou jazyka R</i> .....	57
<b>Kristián Mičko, František Babič</b> <i>Evidencia dočasného parkovania pomocou metód počítačového videnia a umelej inteligencie</i> .....	61
<b>Peter Havran, Roman Cimbala, Maksym Oliinyk</b> <i>Dielektrická spektroskopia transformátorového GTL oleja</i> .....	65
<b>Patrik Hrivniak, Dušan Medveď</b> <i>Využitie open-source platformy na elektroenergetické účely</i> .....	70
<b>Martin Chlebovec, Miloš Drutarovský</b> <i>Bezpečná vzdialená aktualizácia IoT senzora na báze ESP32</i> .....	76
<b>Ján Presada, Dušan Medveď</b> <i>Tepelná analýza priameho kontaktu</i> .....	82
<b>Pavol Silagyi, Dušan Medveď</b> <i>Analýza výroby elektrickej energie pri zbernicovom systéme prehriatej pary</i> .....	88
<b>Peter Havran, Roman Cimbala, Vladimír Kohan</b> <i>Impedance analysis of inhibited mineral oil</i> .....	94
<b>Ján Mihalík</b> <i>Kódovanie ľudskej hlavy v štandarde MPEG-4 SNHC</i> .....	100
<b>Michal Švingál, Dušan Kocur, Jána Fortes</b> <i>UWB-PerLoc-2D3D: softvér na lokalizáciu osôb v 2D a 3D priestore</i> .....	105
<b>Jakub Oravec, Luboš Ovseník, Maroš Lapčák, Norbert Zdravecký</b> <i>Vylepšenie vlastností logistického zobrazenia z hl'adiska obrazových šifrovacích algoritmov</i> .....	110

<b>Jakub Oravec, Luboš Ovseník, Norbert Zdravecký, Maroš Lapčák</b> <i>Parametrizovaný kvantizačný algoritmus pre obrazové šifrovacie algoritmy</i> .....	115
<b>Maroš Lapčák, Norbert Zdravecký, Luboš Ovseník, Jakub Oravec</b> <i>Design of a Backup RF Link for Hybrid FSO/RF System</i> .....	119
<b>Maroš Lapčák, Norbert Zdravecký, Ján Turán, Samuel Andrejčík</b> <i>Monitoring Weather Conditions for Hybrid FSO/RF System Using UAV Devices</i> .....	124
<b>Maroš Lapčák, Norbert Zdravecký, Luboš Ovseník, Jakub Oravec</b> <i>Analysis of the Proposed Horn Antenna as a Backup Line for Hybrid FSO/RF System</i> .....	128
<b>Samuel Andrejčík, Ján Turán, Maroš Lapčák</b> <i>CMS server pre Laboratórium optoelektronických systémov</i> .....	133
<b>Norbert Zdravecký, Luboš Ovseník, Maroš Lapčák, Jakub Oravec</b> <i>Investigation of amplifiers in DWDM systems</i> .....	138
<b>Norbert Zdravecký, Luboš Ovseník, Maroš Lapčák, Jakub Oravec</b> <i>Investigation of modulation in ultra-DWDM systems</i> .....	142
<b>Marek Fedor, Daniela Perduková</b> <i>Dynamická identifikácia vnútorných veličín asynchrónneho motora</i> .....	146
<b>Richard Nezník, Ján Papaj</b> <i>Využitie technológie Blockchain v 6G a MANET sieti</i> .....	154
<b>Filip Mondok, Anton Baláž, Emília Pietriková</b> <i>Web kiosk v OS Linux</i> .....	160
<b>Heidar Khorshidiyeh, Branislav Madoš, Anton Baláž</b> <i>Kompresia multi-dimenzionálnych dát s využitím transformácie ich dimenzionality</i> .....	165
<b>Miroslav Andraš, Dušan Medved'</b> <i>Návrh nekonvenčného zdroja energie na konvenčné účely</i> .....	171
<b>Peter Havran, Roman Cimbala, Miloš Šárpataky</b> <i>Impedance spectroscopy of liquid dielectrics</i> .....	177
<b>Bohuslav Potočňák, Zdeněk Havlice, Lucia Veselovská</b> <i>Application for Truck Entry and Traffic Tracking</i> .....	183
<b>Daniel Slinčák, Zdeněk Havlice, Richard Széplaki</b> <i>System for Processing and Analyzing Test Results</i> .....	189
<b>Jerguš Kačmár, William Steingartner</b> <i>Doc2vec as an expansion to full-text search engine for a specific domain</i> .....	195
<b>Marek Grešák, Dušan Medved'</b> <i>Inteligentné elektromery a ich využitie</i> .....	200
<b>Peter Havran, Roman Cimbala, Luboš Šárpataky</b> <i>Dielectric spectroscopy of transformer paper in the time domain</i> .....	205
<b>Jozef Frivaldský, Marián Mach</b> <i>Porovnanie transformačných operátorov horolezeckého algoritmu v rámci domény TSP</i> .....	210
<b>Peter Focko, Ján Genči</b> <i>Spracovanie EKG záznamov metódami strojového učenia</i> .....	216

<b>Tamás Porteleky, Dušan Kocur, Jána Fortes</b> <i>Softvér v programovacom prostredí LabView určený na sledovanie pohybujúcich sa osôb v reálnom čase</i> .....	222
<b>Marek Roháč, Miloš Drutarovský</b> <i>Bezpečnostné riziko pri generovaní náhodných dát v prostredí virtuálnych strojov s OS Windows</i> .	227
<b>Eva Kupcová, Miloš Drutarovský</b> <i>NTT s konštantnou dobou výpočtu v post-quantovej kryptografii</i> .....	233
<b>Jozef Doci, František Babič, Viera Pavlišková</b> <i>Vizualizácia geopriestorových dát o hovoroch na tiesňovú linku</i> .....	239
<b>Jákob Ildža, Dušan Medved</b> <i>Účinky elektromagnetického poľa na ľudský organizmus</i> .....	244
<b>Miroslav Biñas, Patrik Koscelanský</b> <i>Automation in creation of lists for tests execution in Arena</i> .....	250
<b>Filip Zverec, Ján Papaj</b> <i>Smerovanie a strojové učenie v MANET</i> .....	257
<b>Michaela Šaková, Ján Papaj</b> <i>Bezpečnosť a Blockchain v MANET</i> .....	263
<b>František Jakab, Lucia Michalénková, Dávid Cymbalák, Roman Vápeník</b> <i>Livestreaming Delivery with Enhanced Interactive Functions in NTI in SANET</i> .....	269
<b>František Jakab, Dávid Cymbalák, Miroslav Michalko, Martin Korének</b> <i>Multikamerová distribúcia streamovaného videa z UAV pri sledovaní pohybujúceho sa objektu</i> ..	274
<b>Patrik Schütz, Ján Vaščák</b> <i>Simulation of Activities in a Robotic Swarm</i> .....	279
<b>Maroš Baumgartner, Jozef Juhár, Ján Papaj</b> <i>Výkonová analýza prístupových technológií 5G sietí v simulátore NS-3</i> .....	285
<b>Marek Berka, Ondrej Kainz, Miroslav Michalko, František Jakab</b> <i>Controlling plant growing condition using machine learning</i> .....	291
<b>Jozef Eliáš, Ondrej Kainz, Miroslav Michalko, František Jakab</b> <i>Smart IoT security solution for premises monitoring</i> .....	296
<b>Matúš Huba, Ján Paralič</b> <i>Predikovanie úmrtí na Covid19 v rôznych časových oknách</i> .....	301
<b>Lenka Dobranská, Viera Pavlišková, František Babič</b> <i>Vizualizácia publikačnej činnosti výskumného tímu</i> .....	307
<b>Ján Vaščák, Dušan Herich</b> <i>Cloud Based System for Autonomous Vehicle Navigation</i> .....	312
<b>Miloš Šárpataky, Irida Kolcunová, Ján Zbojovský, Juraj Kurimský, Peter Havran</b> <i>Radiofrequency Shielding Enhancement of Building Materials by Carbon-Based Coating</i> .....	318
<b>Viktor Petro</b> <i>Vybrané metódy bezsnímačového riadenia synchrónneho motora s permanentnými magnetmi</i> ...	327

<b>Martin Hasin, Martin Havrilla</b> <i>Implementácia RSS algoritmu v procese meranie sieťovej prevádzky</i> .....	331
<b>Ján Vaščák</b> <i>Rekurentné fuzzy neurónové siete pri riadení pohonov mobilných robotov</i> .....	337
<b>Gabriela Vašková, Ján Papaj</b> <i>Viacpreskoková mobilné siete MANET a mobilné siete 6. generácie</i> .....	343
<b>Tomáš Maniak, Marek Pavlík</b> <i>Parametre náhradného modelu fotovoltického článku</i> .....	348
<b>Zuzana Pella, Juraj Kanuščák</b> <i>Analýza vplyvu predspracovania dát na vybrané modely analýzy dát</i> .....	352
<b>Tomáš Ondruško, Marek Pavlík</b> <i>Modelovanie účinnosti tienenia elektromagnetického poľa</i> .....	358
<b>Ján Vaščák</b> <i>Návrh inteligentného požiarneho varovného systému s využitím neurónových fuzzy prostriedkov</i> .	362
<b>Peter Girovský, Martin Závaš</b> <i>Riadenie BLDC motora</i> .....	368
<b>Peter Girovský, Stanislav Mačák</b> <i>Návrh algoritmu pre rozpoznávanie obrazu</i> .....	373
<b>Stanislav Ondáš, Miriama Fedorišinová</b> <i>Časové charakteristiky prekrývajúcej sa reči v konverzačných interakciách</i> .....	377
<b>Lukáš Koniar, František Ďurovský</b> <i>Riadenie horizontálneho zásobníka kontinuálnej linky</i> .....	382
<b>Eduard Pynzar, Matúš Pleva</b> <i>Overovanie hodnovernosti biometrického znaku</i> .....	388
<b>Daniel Majerník, Matúš Sulír</b> <i>Audiovisual Live Coding with Audience Participation via Parameter Guessing</i> .....	393
<b>Dávid Bodnár, František Ďurovský</b> <i>Riadenie zásobníka energie elektrického midibusu</i> .....	399
<b>Dávid Gula, Štefan Korečko, Branislav Sobota</b> <i>Prototypovanie riadenia dronov v zdieľanej webovej virtuálnej realite</i> .....	406
<b>Ján Vaščák</b> <i>Návrh systému na reguláciu klímy v skleníku s využitím siete ANFIS</i> .....	412
<b>Daniel Novák, Michal Kolárik</b> <i>Nástroje a knižnice pre analýzu dát v oblasti medicíny</i> .....	417
<b>Maroš Baumgartner, Jozef Juhár, Ján Papaj</b> <i>Simulovanie LTE a 5G prístupových technológií z pohľadu energetickej účinnosti s využitím NS-3</i> .....	422
<b>Ivan Čík, Marián Mach</b> <i>Concept of Concentration Metric for Convolutional Neural Networks</i> .....	428

<b>Branislav Sobota, Štefan Korečko, Matej Masrna</b> <i>Virtuálne prostredie pre podporu vzdelávania hendikepovaných osôb</i> .....	435
<b>Peter Michalik, Ladislav Pomšár, Iveta Zolotová</b> <i>Design Of Methodology for Selection Suitable Analytic Platform in the Area of Data Analysis</i> ..	442
<b>Ján Vaščák</b> <i>Tvorba zabudovaných systémov snímačov pre potreby inteligentného priestoru</i> .....	452
<b>Adrián Marcinek, František Ďurovský</b> <i>Matematický model výkonového multiport meniča</i> .....	458
<b>Tomáš Talian, Štefan Korečko, Branislav Sobota</b> <i>Perspectives of OR/TS2JavaConn Utilization in Automated Rail Control Simulation</i> .....	465
<b>Author's index</b> .....	470

# Electrophysical properties of solid insulation materials

<sup>1</sup>Martin Kolek, <sup>2</sup>Roman Cimbala, <sup>3</sup>Peter Havran

<sup>1</sup> Department of Electric Power Engineering, 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>3</sup> Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical university of Košice, Slovak republic

<sup>1</sup>martin.kolek@student.tuke.sk, <sup>2</sup>roman.cimbala@tuke.sk, <sup>3</sup>peter.havran@tuke.sk

**Abstract** — This publication deals with the investigation of the electrophysical properties of XLPE insulation of a high-voltage cable in the frequency domain. The theoretical part contains a description of electrophysical properties based on polarization mechanisms and the essence of the impedance spectroscopy method. The practical part contains the design of the experiment, analysis of the measured sample, and analysis of the obtained results with a graphical representation of the capacitance of the serial and parallel model of the dielectric and the dissipation factor at frequency.

**Keywords** — dielectric polarization, HV cable, impedance spectroscopy, XLPE insulation

## I. INTRODUCTION

Insulation materials are among the most important parts of electrical engineering and equipment. Without them, it would not be possible to ensure the transmission of electricity through cable and overhead lines. They serve to separate living parts and at the same time prevent direct contact with live parts. Their use is huge, and they are found in cables, transformers, etc. Their properties may be affected by human activity, whether or exceeding their limit values. They play a very important role in high-voltage cables, which are used in substations, mines, or dense construction networks.

In this publication, the insulation system of a high-voltage cable was measured and investigated. The insulating part of the HV cable was made of cross-linked polyethylene, known in short as XLPE. In this XLPE insulation experiment, the voltage dependence and dielectric properties were determined by applying different voltages to the measured sample. One sample of HV cable was measured. The measurements were performed using a non-destructive method of impedance spectroscopy, which plays an important role in the diagnostics of electrical insulation materials. This method measures a small linear dielectric response signal and analyses it to obtain information on the physicochemical properties of the insulation material.

## II. ELECTROPHYSICAL PROPERTIES OF INSULATION MATERIALS AND METHOD FOR THEIR INVESTIGATION

### A. Dielectric polarization

By breaking the symmetry of electric charges, electric dipoles are created in the dielectric, where each dipole has its electric dipole moment. This process is called dielectric polarization. The forces of the external and internal electric field cause the electrically coupled charges to eject from the dielectric only if they are in equilibrium positions. If there are dipole molecules in the substance, they start rotating in the direction of the field. Induced dipoles begin to form from the original equilibrium positions by extending the center of gravity of the opposite charges. The basic property of a dielectric is polarizability, which is the basis of this process. Polarizability can be understood as the orientation of dipole molecules in an electric field. The dissipation factor  $\tan \delta$  and capacitance  $C$  are the main parameters of electrical polarizability and are related to dielectric losses and aging of insulators [1].



The polarization process begins to disappear after the electrical voltage is disconnected with the time constant  $\tau_i$ . The settling time ranges from  $10^{-13}$  to  $10^4$  s. Polarization events in the insulating material have different settling times. For polarization with a settling time greater than  $10^{-6}$  s, the value of  $\tau$  fluctuates around one value (migration polarization). From  $10^{-6}$  to  $10^{-13}$ , the time constants  $\tau$  have a specific value (deformation polarization). The combined polarization spectrum is formed when the time constants of the polarizations overlap, as shown in Fig. 1 [2].

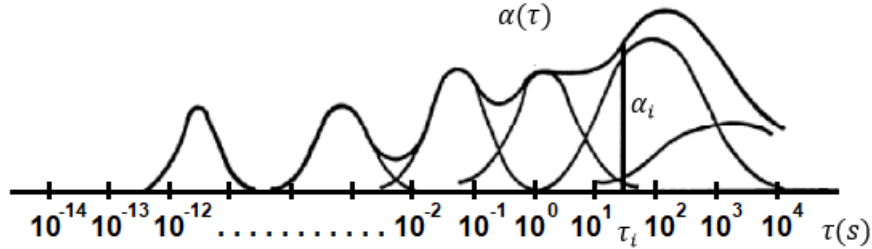


Fig. 1 Polarization spectrum [2]

With one measuring method, we cannot contain the entire polarization spectrum when measuring dielectric quantities. We can measure a maximum of three to four decades, so we must use more measurement methods (infrared spectroscopy, alternating and unidirectional methods). Thanks to this, we obtain information about the entire polarization spectrum. With Maxwell's surrogate scheme, we can express the polarization spectrum in the dielectric. Fig. 2 shows a Maxwell substitute dielectric diagram, where  $C_0$  represents the geometric capacitance,  $R_0$  is the insulation resistance, and the members  $R_i$  and  $C_i$  represent the polarization with stabilization time  $\tau_i$  [2].

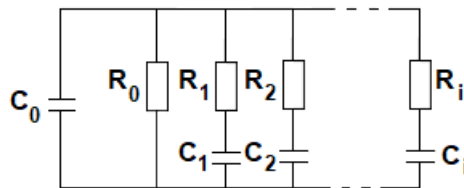


Fig. 2 Maxwell substitute dielectric diagram [2]

### B. Impedance spectroscopy

Each type of insulating material consists at the atomic level of negative and positive charges, which balance each other on microscopic as well as macroscopic scales. As soon as the external electromagnetic field begins to act, polarization events begin to take place in the real material. Due to these polarization changes, it is possible to observe voltage and current changes. Impedance spectroscopy hereinafter referred to as IS, is used to detect responses in material under the influence of an electromagnetic field. Due to degradation, the insulation parameters change, so it is very important to compare them. The future process of degradation is obtained thanks to mathematical-physical models. IS analysis is performed in the time and frequency domain. In the time domain, charging and discharging currents are observed when a DC electric field is applied. In the frequency domain, a dependence on the frequency of complex permittivity, impedance, and parameters such as capacitance, a dissipation factor as a function of temperature and electric field strength, is observed. IS is a method in which there is no damage to the insulating material during the measurement [3].

Measurements can be made using an electrode system that is of the same material and shape as the electrodes. The electrodes can be in the shape of a cylinder, a prism, a hemisphere, or a Rogowski electrode. During the measurement, a current or voltage pulse is applied to the electrodes and its changes are then compared [4][5].

### III. PERFORMANCE OF THE EXPERIMENT OF MEASUREMENT OF ELECTROPHYSICAL PROPERTIES OF XLPE CABLE

In this experiment, the properties of XLPE insulation, which is a part of HV cables, were measured. With XLPE, the mechanical, chemical, thermal, and electrical properties of the material are improved due to the crosslinking of macromolecular fibers. The measured sample of the HV cable consisted of the core, the inner semiconducting layer, XLPE insulation, the outer semiconducting layer, the shielding of the HV cable, the water-blocking foil, and the outer

sheath of PVC. When measuring XLPE insulation, the core of a single-core cable was used as the electrode system and the other electrode was formed by the shielding of the HV cable. The sample examined was 1.2 m long. A measurement method in the frequency domain of impedance spectroscopy was used. The Megger IDAX 300 measuring instrument was used in the frequency domain, where the values of the dissipation factor  $\tan \delta$  and capacitance  $C$  were measured and evaluated in a serial and parallel measuring model using alternating voltage. The measurement of the dissipation factor is very important because it determines the condition of the insulation material. Voltages of 1, 2, 10, and 100 V in the frequency band from 1 mHz to 10 kHz were used in the measurement. To achieve a steady-state value, three measurements were taken for each voltage value. All measurements were performed in a shielded EMC chamber due to electromagnetic interference (Fig. 3).

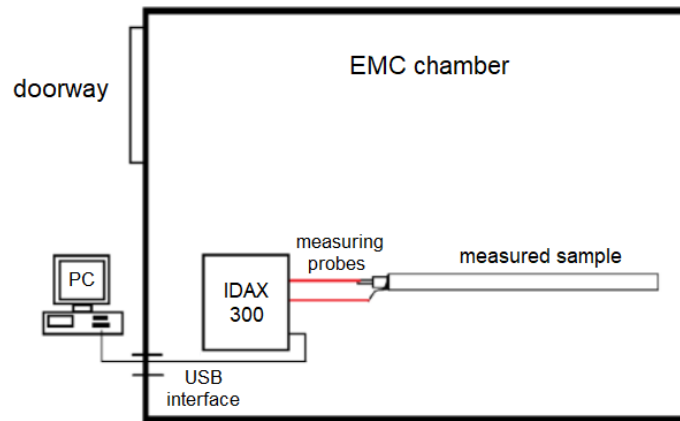


Fig. 3 Connection of measuring workplace

#### IV. ANALYSIS OF THE RESULTS OBTAINED

In this part, there is an analysis of measured values in the frequency domain when applying an alternating voltage of 1, 2, 10, and 100 V. The dissipation factor and capacitance were measured in a parallel and series measuring model.

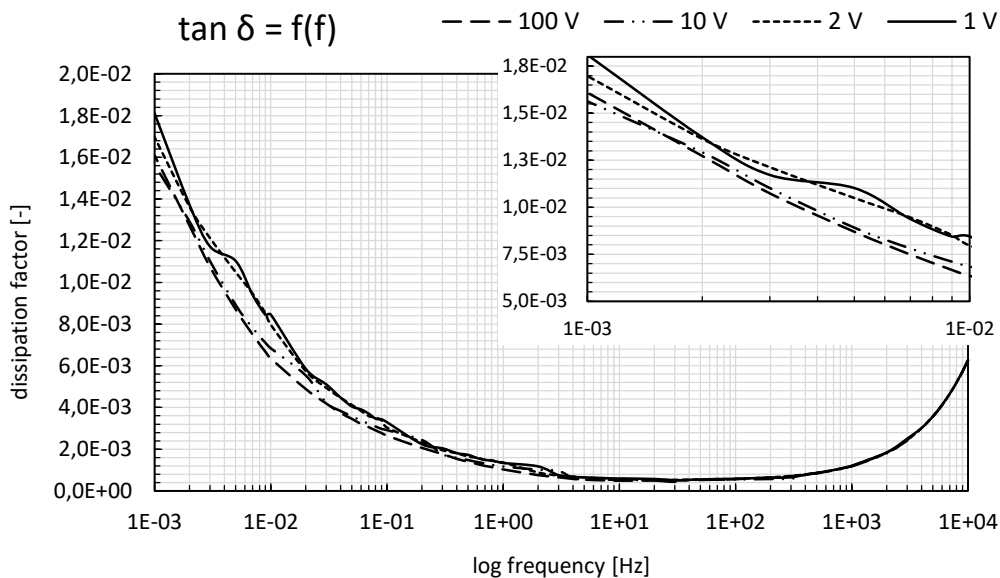


Fig. 4 Comparison of dissipation factor values when applying different voltages

For a better comparison of the measured values, the graphical dependence of the loss factor on the frequency in Fig. 4 was recorded by the frequency axis on a logarithmic scale. The dissipation factor determines the quality of the insulation. The value of the dissipation factor at all voltage values decreases with increasing frequency. The highest value of the dissipation factor was measured at the value of applied voltage 1 V. At very low frequencies (1 mHz) the lowest value of dissipation factor was measured at the applied voltage 10 V, with the lowest dissipation factor having a voltage loss of 100 V. The displayed curve corresponds to the theoretical knowledge. The dissipation factor measurement aimed to determine whether the

value of the applied voltage affects the value of the dissipation factor. It can be seen from the graphical dependencies that the value of the dissipation factor is very little affected depending on the applied voltage. The dissipation factor is most affected in the frequency range from 1 mHz to 1 Hz. In the medium and higher frequencies, the dissipation factor is almost the same for all applied voltages. At very low frequencies, the applied AC voltage may appear as a DC voltage, where the time change of one-half period is too long and therefore only the conductivity component is applied here. In the region of medium frequencies, the conductivity component of the dissipation factor is stabilized and with increasing frequency, a polarization event begins to develop with a certain relaxation time, which is not known due to the insufficient frequency range of the measuring instrument.

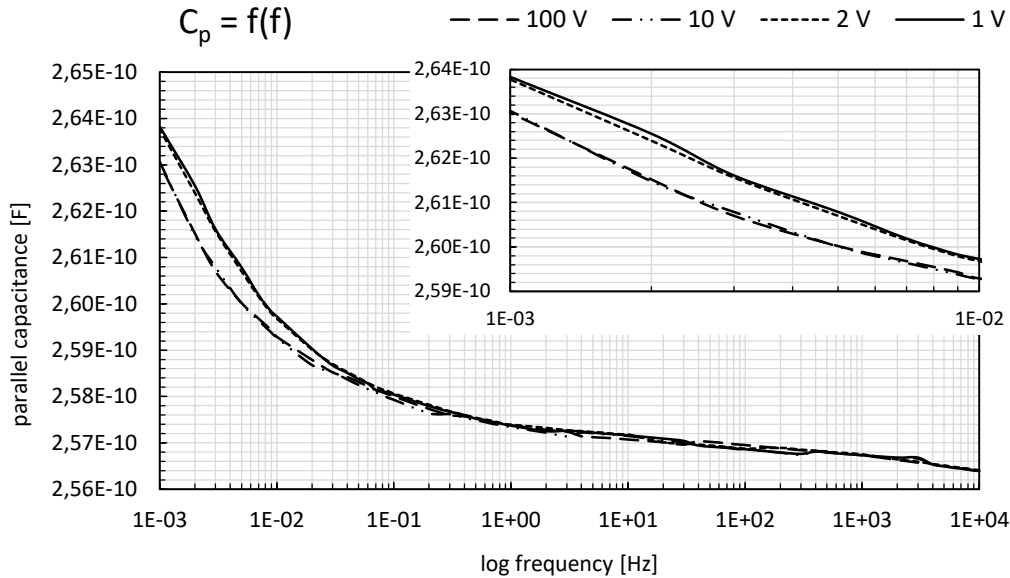


Fig. 5 Comparison of capacitance values of a parallel model when applying different voltages

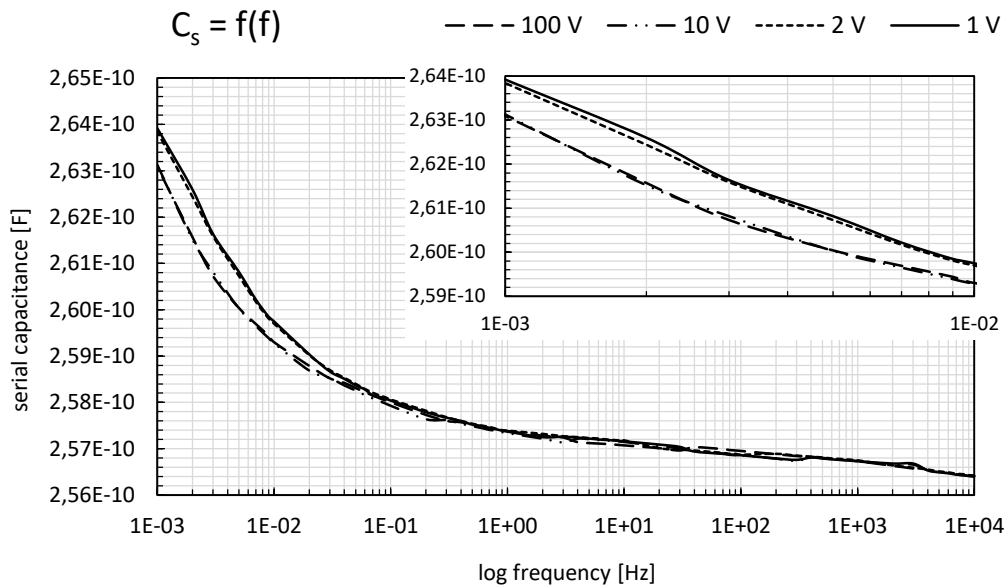


Fig. 6 Comparison of capacitance values of a serial model when applying different voltages

In Fig. 5, the capacitance was measured in the parallel measurement model, which means that in the replacement measurement model, the capacitor is connected in parallel to the resistor. Fig. 6 is in comparison with Fig. 5 measured capacitance in a series measuring model, which means that the capacitor is connected in series with a resistor. With both dependences in Fig. 5 and Fig. 6 capacitance values when using a serial or parallel measurement model do not differ. Also, the value of the measuring voltage does not partially affect the size of the capacitance in both cases. Conversely, the size of the measured sample capacitance depends on the size of the AC voltage frequency. The dependencies have a decreasing character, which means that the higher the frequency value, the lower the capacitance value. The dependence on the voltage

used was most pronounced in the frequency range from 1 mHz to 0.1 Hz. It was found that a higher value of capacitance at low frequencies in the measured sample was achieved at lower voltages at 1 and 2 V. It can therefore be said that in this frequency band the capacitance decreases with increasing voltage. From 0.1 Hz, the capacitance has a slightly decreasing character approaching the linear character of the curve and the value of the capacitance almost does not change depending on the magnitude of the voltage. The difference in capacitance size is around 0.031%. These courses can be combined with theoretical knowledge.

## V. CONCLUSION

This paper aimed to determine the rate of change of the dissipation factor and the capacitance (serial and parallel model) of XLPE insulation in the frequency band at different applied voltages. It can be seen from the graphical dependencies that the value of the dissipation factor is very little affected depending on the applied voltage. The dissipation factor is most affected in the frequency range from 1 mHz to 1 Hz when the dissipation factor decreases with increasing voltage. In the region of medium and higher frequencies, with the developing polarization mechanism, the dissipation factor is the same for all applied voltages. Capacitance values do not differ when using a serial or parallel model. Also, the voltage value does not partially affect the size of the capacitance in both models. The dependence on the voltage used manifested itself in the frequency range from 1 mHz to 0.1 Hz. In this frequency band, the capacitance decreases with increasing voltage. From 0.1 Hz to 10 kHz, the capacitance hardly changes depending on the voltage.

## ACKNOWLEDGMENT

The work was supported by the Ministry of Education, Science, Research and Sport within the project VEGA 2/0011/20 and 1/0154/21 and the Slovak Agency for Research and Development based on contract no. APVV-15-0438, APVV-17-0372, and APVV-18-0160.

## REFERENCES

- [1] A. Pietriková, J. Banský, „Základy inžinierstva materiálov“, Technical university of Košice, Letná 9, Košice. 2007, 278 pp., ISBN 80-8086-001-7.
- [2] I. Kolcunová, „Dielektrické veličiny“, Online: <https://moodle.tuke.sk/moodle35/mod/folder/view.php?id=20974>.
- [3] J. R. Macdonald, „Impedance spectroscopy“, In. *Annals of Biomedical Engineering*. 1991, 20, 2 pp. ISSN 1573-9686.
- [4] E. Barsoukov, J. R. Macdonald, „Impedance spectroscopy: Theory experiment and applications“, New Jersey: John & Sons, Inc. 2005, 608 pp. ISBN 0-471-64749-7.
- [5] F. Vadim, „Impedance spectroscopy: Applications to Electrochemical and Dielectric Phenomena“, New Jersey: John Wiley & Sons, Inc. 2012, 356 pp. ISBN 978-0-470-62778-5.

# Návrh automatizovaného systému pre tvrdé prepínanie hybridnej FSO/RF linky na báze strojového učenia

<sup>1</sup>Luboš OVSEŇÍK, <sup>2</sup>Jakub ORAVEC, <sup>3</sup>Maroš LAPČÁK, <sup>4</sup>Norbert ZDRAVECKÝ

<sup>1,2,3,4</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>lubos.ovsenik@tuke.sk, <sup>2</sup>jakub.oravec@tuke.sk, <sup>3</sup>maros.lapcak@tuke.sk,

<sup>4</sup>norbert.zdravecky@tuke.sk

**Abstrakt** — Táto práca sa zaoberá dostupnosťou bezvláknových optických FSO systémov. Pre dosiahnutie 99,999 % dostupnosti bolo nutné vytvorenie hybridného FSO/RF systému z dôvodu nepriaznivého vplyvu poveternostných podmienok (sneh, hmla, dážď, prach, smog, dym) na prenosové vlastnosti FSO systému. V súčasnosti je dostupné „mäkké“ a „tvrdé“ prepínanie záložnej RF linky v prípade nepriaznivých prenosových podmienok. V tejto práci sme sa venovali tvrdému prepínaniu pri poklese prijatého optického výkonu (RSSI) pod prahovú hranicu -42 dBm, ktorú sme získali dlhodobým výskumom na experimentálnom FSO/RF systéme umiestnenom v areáli TUKE.

**KLúčové slová** — FSO systémy, hybridné FSO/RF systémy, prepínanie linky, RF systémy, RSSI

## I. ÚVOD

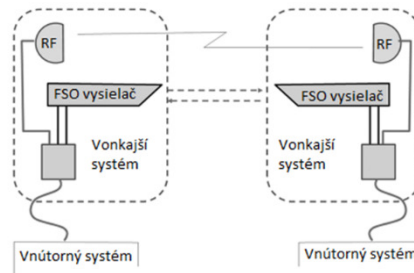
Dopyt po veľkej šírke pásma v metropolitných sieťach s krátkym časom prístupu sa zvyšuje. Okrem toho požiadavky pružnosti a nákladovej efektívnosti poskytovania služieb spôsobujú určitú nerovnováhu. Tento problém sa označuje ako „prekážka na poslednom kilometri“ (z angl. last mile bottleneck).

Prvá najpoužívanejšia možnosť riešenia nedostatku šírky pásma je využitie optického vlákna ako prenosového média. Vlákno je bezpochyby najspoľahlivejším prostriedkom optickej komunikácie, avšak náklady na vytvorenie prenosovej infraštruktúry (uloženie optického vlákna) a čas uvedenia takého systému na trh sú jej najzávažnejšími nevýhodami. Druhou možnosťou je komunikácia cez rádiovú frekvenciu (RF). Táto technológia je vyspelá a v dnešnej dobe často nasadzovaná (GSM, WiFi, WiMAX). Siete založené na rádiových sieťach vyžadujú obrovské investície na získanie licencií na frekvenčné spektrum, avšak nemôžu sa porovnávať s prenosovou kapacitou optických vlákien. Treťou alternatívou sú všetky technológie na báze metalickej prenosovej linky, ako sú káblové modemy a DSL (Digital Subscriber Line). Napriek tomu, že „medená“ infraštruktúra je k dispozícii v širokom meradle, problémom je škálovateľnosť šírky pásma. Štvrtou alternatívou je komunikácia voľným priestorom FSO (Free Space Optics). FSO predstavuje optimálne riešenie z hľadiska technológie, škálovateľnosti šírky pásma, rýchlosti nasadenia a efektívnosti nákladov. Optická bezdrôtová komunikácia ponúka alternatívu k rádio-frekvenčnej komunikácii aj v prípadoch, kde sú požadované vysoko výkonové linky. Hlavná podmienka pri využívaní FSO spočíva v potrebe priamej viditeľnosti medzi vysielačom a prijímačom.

Hybridná FSO/RF bezdrôtová komunikácia voľným priestorom (FSO) a komunikácia cez rádiové frekvencie (RF) je spôsob, ako zabezpečiť spoľahlivú komunikáciu pre kritickú prevádzku v reálnom čase, aj keď počasie, ako je hmla, dážď, sneh či znečistenie vzduchu, má veľký vplyv na dostupnosť takéhoto systému. Keďže FSO komunikácia je ovplyvňovaná poveternostnými podmienkami pozdĺž prenosovej cesty komunikačnej linky, je potrebné navrhnúť a vytvoriť spôsob prepínania medzi primárnou FSO a záložnou RF linkou. Pre potreby takéhoto samo-rozhodovacieho prenosového systému je nutný systém pre zber poveternostných dát spolu s prijatým optickým výkonom (RSSI). Získané dáta sa následne ukládajú do databázy a vhodnými metódami strojového učenia môžeme predikovať prijatý optický výkon (RSSI) o určitú časovú jednotku dopredu.

## II. HYBRIDNÉ FSO/RF SYSTÉMY

Očakáva sa, že dostupnosť spojenia FSO bude lepšia ako 99,999 %, no v niektorých oblastiach sa častejšie vyskytuje hustá hmla, krupobitie, sneh, dážď, piesočné búrky, atď. Keď sa hustá hmla ustáli v oblasti, na niekoľko hodín, zničí predpísanú očakávanú dostupnosť služieb. Aby služba bola k dispozícii potrebných 99,999 %, je nutná záložná komunikačná stratégia; pre pozemné aplikácie táto stratégia zahŕňa symbiotickú hybridnú technológiu, ktorá pozostáva z primárneho spojenia FSO pracujúceho paralelne so sekundárnym širokopásmovým rádiovým (RF) spojením; ako náhle sa spojenie FSO stane nefunkčným alebo jeho výkon klesne pod očakávanú úroveň, RF linka sa stáva funkčnou na pokračovanie nepretržitej služby (Obr. 1), pričom RF systém využíva anténu s vysokým ziskom a s vysokou smerovosťou [1].



Obr. 1 Hybridný FSO/RF prenosový systém.

Výkon FSO komunikácie je drasticky ovplyvnený poveternosťnými podmienkami a turbulenciami v atmosfére. To môže viesť k zlyhaniu spojenia alebo k zníženiu prenášaného optického výkonu. Preto, aby sa zvýšila spoľahlivosť a zlepšila dostupnosť spojenia, je rozumné spárovať systém FSO so záložným a spoľahlivým RF systémom. Takéto systémy sa nazývajú hybridné RF/FSO a sú schopné poskytnúť vysokú dostupnosť spojenia aj za nepriaznivých poveternosťných podmienok [1]. Hlavnou príčinou degradácie signálu vo vysielaní RF je dážď (pretože nosná vlnová dĺžka je porovnateľná s veľkosťou dažďovej kvapky) a komunikácia FSO je najmarkantnejšie ovplyvňovaná hmlou. Takže celková dostupnosť systému môže byť vylepšená použitím linky s nízkou rýchlosťou prenosu dát ako zálohy, keď je prepojenie FSO nefunkčné. Zistilo sa, že spojenie FSO poskytuje nízku dostupnosť v podmienkach nízkej oblačnosti v dôsledku útlmu častíc mraku a časovej disperzie. Značné zlepšenie sa však pozorovalo, keď sa použilo hybridné spojenie FSO/RF, pretože RF signály sú imúnne voči interferencii s oblakom. Bežný prístup hybridných FSO/RF spôsobuje neefektívne využitie šírky pásma kanálu [2]. Taktiež nepretržité prepínanie medzi systémom RF a systémom FSO by mohlo znefunkčnúť celý systém. Preto nový prístup poskytuje schému adaptačného kódovania symbolovej rýchlosti, v ktorej sú súčasne aktívne aj RF a FSO subsystemy a šetria tak šírku pásma kanálu. Kódovanie hybridného kanála je tiež schopné využívať obe prepojenia kombináciou nehomologických kódov a kódov prispôbujúcich rýchlosť, kde sa ich kódové rýchlosti menia podľa podmienok kanálov.

Hybridné prepojenie FSO/RF poskytuje skvelú aplikáciu v mobilných sieťach ad-hoc (MANET) [3]. Rekonfigurovateľné sieťové prostredie môže byť vytvorené v MANET kombináciou technológie bezdrôtovej senzorovej siete (WSN) a mobilnej robotiky. Výkonnosť tejto siete je však obmedzená výkonom na uzol, ktorý poskytuje komunikácia založená na RF. Preto kombinácia RF a FSO poskytuje obrovský nárast prenosu MANETov na každý uzol. Bezdrôtová sieť RF predstavuje silné obmedzenie kapacity a výkonnosti v dôsledku rastúceho rozvoja komunikačných technológií [2]. S narastajúcim počtom používateľov sa zvyšujú šance na rušenie zo susedných uzlov a obmedzujú výkonnosť RF systému. Systém FSO je na druhej strane veľmi smerový a má veľmi úzku divergenciu lúča. To robí systém FSO odolný voči akémukoľvek rušeniu. Preto kombinácia FSO a RF môže pomôcť pri riešení problému nedostatku kapacity v sieťach RF. Keď sa optický signál šíri cez atmosférický kanál, naráža na zmenu intenzity signálu v dôsledku rôznych nepredvídateľných faktorov prostredia, ako je hmla, dážď, sneh, atď. [3].

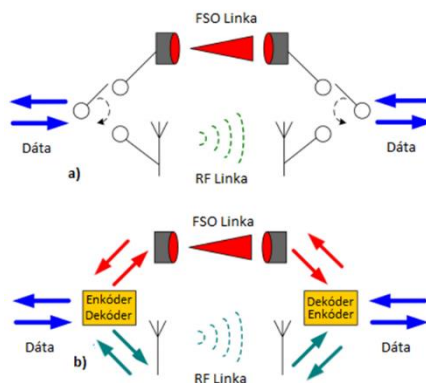
### A. Metódy prepínania v hybridných FSO/RF systémoch

Vzhľadom na odlišné nosné frekvencie majú FSO a RF linky rozdielne kanály. Krátke vlnové dĺžky použité v spojeniach FSO sú vážne ovplyvnené rozptýlením v dôsledku hmly a atmosférickej scintilácie, zatiaľ čo RF sú obzvlášť citlivé na dažďový rozptyl. Použitie oboch systémov na prekonanie týchto problémov môže viesť k zvýšeniu spoľahlivosti a rýchlosti [4].

Pri konfigurácii tvrdého prepínania FSO linky vysielateľ a prijímač spoločne vyberajú buď FSO

alebo RF kanál pre dátový prenos. Vysokorychlostné spojenie FSO sa volí len vtedy, ak podmienky kanálu umožňujú spoľahlivú komunikáciu, inak sú všetky údaje odosielané cez RF kanál. Vysielač a prijímač musia byť koordinované prostredníctvom spätnej väzby na výber správneho kanálu pre prenos. Kľúčovou nevýhodou tvrdého prepínania je to, že kedykoľvek je jedno spojenie nečinné. V praxi môže scintilačné pôsobenie alebo strata v citlivom kanáli FSO spôsobiť jeho „zriedkavý výber“.

Po vybratí RF spojenia sa stratí kapacita kanálu spojenia FSO. Obr. 2 predstavuje dve možné konfigurácie hybridného spojenia FSO/RF [5].



Obr. 2 a) Tvrdé, b) Mäkké prepínanie hybridného FSO/RF systému.

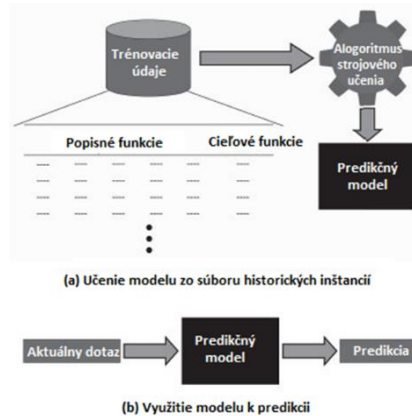
Myšlienka na prekonanie nevýhodného prepínania je koordinovať prenos dát v oboch spojeniach pomocou kódovania kanálov. Dáta sú zakódované jedným LDPC kódom s časťou kódového slova rozdeleného na FSO a RF spojenia a rýchlosť je nastavená podľa okamžitých podmienok kanálov. Aj keď sa táto technika zlepšuje v dôsledku tvrdého prepínania, musí byť na vysielači a prijímači známe podmienky kanálov a vyžaduje sa komplexné jemné dekódovanie, ktoré je pri dátových rýchlostiach FSO ťažké. Preto sa uvažuje o využívaní kódovania Raptor pre hybridné linky FSO/RF [6]. Toto kódovanie nevyžaduje znalosť kanálov na vysielači a je schopné prispôbiť svoju frekvenciu podmienkam kanálu len s jednou bitovou spätnou väzbou na jednu správu. V schéme kódovania Raptor, v ktorej sú bity prenášané na FSO a RF spojoch náhodné lineárne kombinácie bitov správ, sa použije dobre známy kód a predpokladá sa, že vysielač a prijímač majú informácie o prepojení pre každý bit. Navyše je potrebná mäkká iteračná detekcia v prijímači a nepriaznivý vplyv rôznych atmosférických podmienok sa neberie do úvahy [7].

Prijatý optický výkon RSSI (Received Signal Strength Indicator) je podstatný ukazovateľ pre tvrdé prepínanie FSO/RF systému na záložnú linku. V dnešnej dobe sa FSO/RF systém prepne na záložnú linku pri detegovaní nízkej hranice signálu na strane prijímača. Pritom je nutné zabezpečiť informáciu o prepnutí ako na strane prijímača, tak aj na strane vysielača. Pre opätovné prepnutie systému zo záložnej linky na optickú, musí byť hodnota prijatého optického signálu nad hraničnou kritickou hodnotou určitý čas. Nakoľko RF linky dosahujú menšiu prenosovú rýchlosť je vhodné vytvoriť dostupnú predikciu parametra prijatého optického výkonu RSSI. Táto predikcia by hlavne slúžila na určenie potrebného časového intervalu pre komunikáciu cez záložnú RF linku v rámci hybridného FSO/RF systému. Vhodná predikcia prijatého optického signálu a následná redukcia doby komunikácie cez záložnú linku zvyšuje efektivitu využívania vysokorychlostnej prenosovej linky. Avšak je nutné udržiavať nepretržitú komunikáciu v rámci dvoch uzlov hybridnej FSO/RF linky. Lúč posielaný z optickej hlavice FSO je postupne atmosférickým prenosovým kanálom utlmovaný, pričom tento útlm závisí od veľkosti vplyvu poveternostných podmienok. Optický výkon, ktorý je reálne prijatý v danom čase môže taktiež slúžiť ako parameter interpretujúci parametre popisujúce poveternostné podmienky.

### B. Prepínanie linky pomocou strojového učenia

Strojové učenie je definované ako automatizovaný proces, ktorý vyberá vzory z dát. Na vytvorenie modelov používaných v aplikáciách prediktívnej analýzy dát používame tzv. „počítačové učenie pod dohľadom“. Techniky technického dozoru nad strojom sa automaticky naučia model vzťahu medzi súborom popisných funkcií a cieľovou funkciou na základe súboru historických príkladov alebo inštancií. Tento model potom môžeme použiť na predpovedanie nových prípadov [8]. Tieto dva samostatné kroky sú znázornené na Obr. 3.

Algoritmy strojového učenia vyberajú z množiny predikčných modelov taký model, ktorý najlepšie zachytáva vzťah medzi popisnými znakmi a cieľovou funkciou v množine údajov. Zrejme kritériá pre riadenie tohto vyhľadávania sú hľadania modelov, ktoré sú v súlade s údajmi.



Obr. 3 Strojové učenie s dohľadom.

Existujú prinajmenšom dva dôvody, prečo práve hľadanie konzistentných modelov nie je dostatočné na to, aby naučili užitočné predikčné modely. Po prvé, pokiaľ ide o veľké množiny údajov je pravdepodobné, že v údajoch bude šum a predikčné modely, ktoré sú v súlade so zašumeným údajom, budú robiť nesprávne predpovede. Po druhé, vo veľkej väčšine využití strojového učenia tréningový súbor predstavuje len malú vzorku možného súboru inštancií v doméne. V dôsledku toho je strojové učenie „zle položeným problémom“. Zle položený problém je problém, pre ktorý nie je možné určiť jedinečné riešenie pomocou iba dostupných informácií.

Rôzne algoritmy strojového učenia zakódujú rôzne predpojatosti. Pretože algoritmus strojového učenia kóduje indukčnú predpojatosť, môže indukovať modely, ktoré sa zovšeobecňujú za inštancie v tréningovom súbore údajov. Nevhodná indukčná predpojatosť však môže viesť k chybám. Ukázalo sa [9], že neexistuje žiadna konkrétna indukčná tendencia, ktorá by bola v priemere najlepšia na použitie. Taktiež vo všeobecnosti neexistuje žiadny spôsob, ako vedieť pre danú prediktívnu úlohu, ktorá indukčná predpojatosť bude fungovať najlepšie. Schopnosť vybrať vhodný algoritmus strojového učenia (a teda aj indukčnú predpojatosť), ktorý sa má použiť pre danú prediktívnu úlohu, je jednou zo základných zručností, ktoré musí analytik údajov vyvinúť.

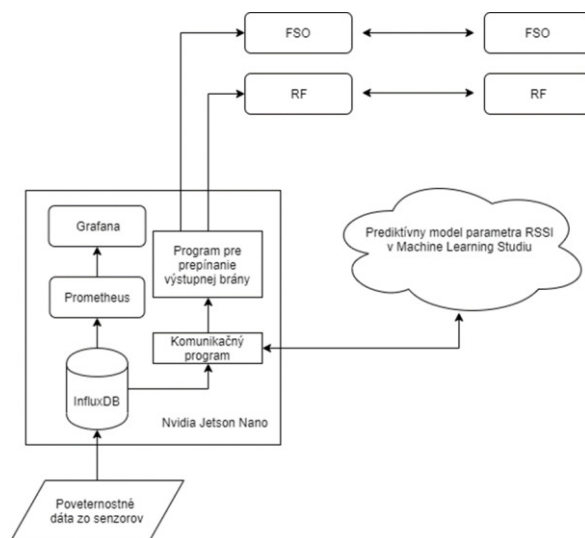
Existujú dva druhy chýb, ktoré môžu viesť k nevhodnej indukčnej predpojatosti: overfitting a underfitting. Overfitting nastáva, keď model predpovede vybraný algoritmom je príliš zjednodušený na to, aby reprezentoval základný vzťah v množine údajov medzi popisnými znakmi a cieľovou funkciou. Underfitting sa naopak vyskytuje vtedy, keď model predpovede zvolený algoritmom je taký komplexný, že model je príliš úzko spojený s dátovým súborom a stáva sa citlivým na šum v údajoch.

### III. NÁVRH AUTOMATIZOVANÉHO UČIACEHO SA SYSTÉMU PRE TVRDÉ PREPÍNANIE

Pre potreby automatizovaného systému, ktorý bude vykonávať predikciu parametra RSSI, bol vytvorený automatizovaný systém zložený z programu pre odosielanie a prijímanie dát v rámci senzorickej stanice, ktorú riadi mini počítač Nvidia Jetson Nano. Ďalšou súčasťou automatizovaného systému je prediktívna webová služba, ktorá pomocou vstupného a výstupného modulu webovej služby predikuje cez zvolený natrénovaný model strojového učenia výstupný parameter RSSI (Obr. 4). Na základe parametra RSSI je následne zvolená výstupná brána pre komunikáciu hybridného FSO/RF systému. Ak parameter RSSI klesne pod kritickú hodnotu  $-42$  dBm, automaticky sa prepne na záložnú RF linku, ak minipočítač získa z prediktívneho modelu hodnotu parametra RSSI, ktorá je lepšia ako kritická, prepne komunikačný kanál späť na primárnu FSO linku s vyššou prenosovou rýchlosťou.

Klasická verzia Microsoft Azure Machine Learning Studio umožňuje nasadiť riešenie predikcie parametra RSSI ako webovú službu. Webové služby Machine Learning Studio poskytujú rozhranie medzi aplikáciou pracujúcou na minipočítači Nvidia Jetson Nano a zvoleným prediktívnym modelom pracujúcim v cloude. Aplikácia môže komunikovať s predikčným modelom v reálnom čase. Volanie webovej služby v Machine Learning Studiu vráti výsledky predikcie späť do aplikácie pracujúcej na minipočítači Nvidia Jetson Nano. Ak chceme využívať predikčný model vytvorený v cloude ako webovú službu, potrebujeme na to kľúč API (z angl. Application Programming Interface), ktorý bol vytvorený spolu s webovou službou.



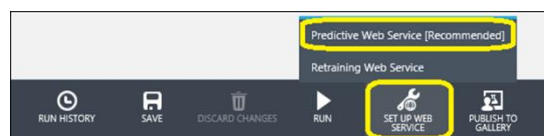


Obr. 4 Blokový diagram automatizovaného učiaceho sa systému pre tvrdé prepínanie hybridného FSO/RF systému.

Webová služba v Machine Learning Studiu je založená na architektúre rozhraní Representational State Transfer (REST), ktorá je veľmi populárnou voľbou architektúry pre projekty webového programovania. Azure Machine Learning Studio má dva typy webových služieb:

1. Služba žiadosti o odpoveď, ide o službu s nízkou latenciou, je vysoko škálovateľná a posiela jeden dátový záznam.
2. Dávková exekučná služba, ide o asynchrónnu službu, ktorá posiela dávku dátových záznamov.

Po natrénovaní zvoleného modelu je možné previesť natrénovaný experiment na prediktívny experiment a získať tak nové údaje. To je v našom prípade konverziou na prediktívny experiment natrénovaný model pripravený v podobe webovej služby (Obr. 5). Používatelia webovej služby môžu poslať vstupné údaje do predikčného modelu a predikčný model odošle späť výsledky predikcie.



Obr. 5 Príklad konverzie natrénovaného modelu na prediktívny model.

Proces premeny tréningového experimentu na prediktívny experiment zahŕňa tri kroky:

1. Výmenu modulov algoritmu strojového učenia za trénovaný model.
2. Odstránenie modulov, ktoré sú potrebné na učenie modelu, ale nie sú potrebné akonáhle je model natrénovaný.
3. Definovanie, ako bude model prijímať údaje z mini počítača Nvidia Jetson Nano a aké údaje sa vrátia späť do mini počítača.

V niektorých prípadoch možno umožniť používateľovi webovej služby zmeniť správanie modulov pri prístupe k službe. Toto umožňujú parametre webovej služby. Bežným príkladom je nastavenie modulu Import údajov, aby používateľ nasadenej webovej služby mohol pri prístupe k webovej službe určiť iný zdroj údajov. Jednou z ďalších možností je nakonfigurovanie modulu Export dát tak, aby bolo možné zadať iný cieľ. Definovaním parametrov webovej služby a priradením ich k jednému alebo viacerým parametrom modulu môžeme určiť, či sú povinné alebo voliteľné. Používateľ webovej služby poskytuje hodnoty týchto parametrov pri prístupe k službe a modul sa zodpovedajúcim spôsobom upraví.

K webovej službe Machine Learning je možné sa pripojiť pomocou ľubovoľného programovacieho jazyka, ktorý podporuje požiadavku a odpoveď HyperText Transfer Protokolu (HTTP). Pre pripojenie k webovej službe Machine Learning, sme použili knižnicu urllib.request pre Python 3.7. Službu Machine Learning sme autentifikovali pomocou kľúča API. Príklad komunikačného programu v jazyku Python 3.7, medzi mini počítačom Nvidia Jetson Nano a Machine Learning Studiom môže vyzeráť nasledovne:

```
import urllib.request #knížnica pre otvorenie ulr
import urllib.error # knížnica pre výnimky vyvolané urllib.request

#vstupné dáta
data = {
    "Inputs": {
        "input1":
            [
                {
                    'temperature' : "26",
                    'humidity' : "30",
                    'wind_speed' : "13",
                    'air_presure' : "980",
                    'visibility' : "4000",
                    'dust_particles' : "35",
                }
            ],
    },
    "GlobalParameters": {}
}

body = str.encode(json.dumps(data)) #metóda na prevod slovníkového objektu pythonu do dátového
formátu reťazca JSON

url = '<api uri aplikácie>'
api_key = '<api kľúč aplikácie>'
headers = {'Content-Type': 'application/json', 'Authorization': ('prediction '+ api_key)} #autorizácia
do Azure

req = urllib.request.Request(url, body, headers)

try:
    response = urllib.request.urlopen(req) #funkcia pre získanie dát
    result = response.read() #čítanie dát
except urllib.error.HTTPError, error:
    print("The request failed with status code: " + str(error.code)) #zachytenie výnimky

    print(error.info())
    print(json.loads(error.read()))
```

#### IV. ZÁVER

Pre vytvorenie automatizovaného systému pre tvrdé prepínanie hybridnej FSO/RF linky na báze strojového učenia bol využitý cloudový systém Microsoft Azure a Machine Learning Studio. Vytvorením prediktívneho modelu ako webové služby vznikla možnosť rýchlej zmeny algoritmu strojového učenia pre tvrdé prepínanie linky. Ako predikčný model bola zvolená metóda Two-Class Decision Jungle, keďže získala najlepšiu percentuálnu úspešnosť predikcie. Porovnaním kritickej hranice -42 dBm, pod ktorou nie je možný prenos cez FSO linku s výstupnou hodnotou z prediktívneho modelu pracujúceho v cloude, sme získali plne automatický systém prepínania hybridnej FSO/RF linky. Analýza údajov je založená na princípe hľadania určitých vzorov, ktoré zdanlivo nesúvisia a náhodne sa v čase menia. Tento spôsob je aplikovaný na pravidelne merané parametre, ktoré slúžia ako vstupné parametre pre natréňovanie predikčných modelov. Na základe analýzy skúmaných predikčných metód je možné aplikovať toto riešenie na navrhovaný systém tvrdého prepínania FSO/RF linky.

#### POĎAKOVANIE

Táto práca bola podporená výskumnými projektami VEGA 1/0584/20 a FEI-2021-71.

#### LITERATÚRA

- [1] L. B. Stotts, at al., "Progress towards reliable free-space optical networks", in *Military Communications Conference MILCOM 2011*, IEEE, 2011, pp. 1720-1726.
- [2] D. W. Young, at al., "Demonstration of high data rate wavelength division multiplexed transmission over a 150 km free space optical link networks", in *Military Communications Conference MILCOM 2017*, IEEE, 2017, pp. 1-6.
- [3] L. C. Andrews, at al., "Hybrid Optical/Radio Frequency (RF) Communications", in *Advanced Free Space Optics (FSO)*, Springer, New York, NY, 2015, pp. 295-342.
- [4] W. Zhang, et al., "Soft-switching hybrid FSO/RF links using short-length raptor codes: design and implementation", *IEEE Journal on Selected Areas in Communications*, 27(9), 2009, pp. 1698-1708.
- [5] A. Abdulhussein, et al., "Rateless coding for hybrid free-space optical and radio-frequency communication", *IEEE Transactions on Wireless Communications*, 9(3), 2010, pp. 907-913.
- [6] D. Killinger, "Free space optics for laser communication through the air", *Optics and Photonics News*, 2002, pp. 36-42.
- [7] A. Eslami, et al., "Hybrid channel codes for efficient FSO/RF communication systems", *IEEE Transactions on Communications*, 58(10), 2010, pp. 2926 - 2938.
- [8] D. E. Goldberg, et al., "Genetic algorithms and machine learning", *Machine Learning*, 1988, pp. 95-99.
- [9] M. I. Jordan, et al., "Machine learning: Trends, perspectives, and prospects", *Science*, 2015, pp. 255-260.

# Návrh 32-kanálového Ultra-DWDM systému pre potlačenie vplyvu nelineárnych javov SPM, XPM a FWM

<sup>1</sup>Luboš OVSENÍK, <sup>2</sup>Jakub ORAVEC, <sup>3</sup>Norbert ZDRAVECKÝ, <sup>4</sup>Maroš LAPČÁK

<sup>1,2,3,4</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>lubos.ovsenik@tuke.sk, <sup>2</sup>jakub.oravec@tuke.sk, <sup>3</sup>norbert.zdravecky@tuke.sk,  
<sup>4</sup>maros.lapcak@tuke.sk

**Abstrakt** — Predmetom tohto príspevku je návrh a analýza 32-kanálového Ultra-DWDM systému s adaptáciou rôznych vysokospektrálnych modulácií, ako nástrojov na regulovanie nežiaduceho vplyvu lineárnych a nelineárnych javov, ktoré sú hlavným faktorom limitujúcim kapacitu a prenosovú vzdialenosť optického komunikačného systému.

**Kľúčové slová** — DWDM, nelineárne javy, optické modulácie, ultra-DWDM

## I. ÚVOD

Exponenciálne rastúci dopyt po prenosovej kapacite vo vysokorýchlostných prenosových systémoch na dlhé vzdialenosti prinútil optickú vlákňovú komunikačnú technológiu, aby si ako životaschopné riešenie osvojila systém multiplexovania s hustým vlnovým delením DWDM (Dense Wavelength Division Multiplexing). Pre maximalizáciu výkonu takýchto optických sietí sa kritickou oblasťou výskumu stali vhodné stratégie návrhu komunikačného systému na optimalizáciu parametrov súvisiacich s prenosom dát v rámci vlastností kanála. Kapacita optického kanála je zvyčajne ovplyvnená tvarom signálu, disperznými a nelineárnymi charakteristikami optického vlákna. Preto, aby sa prenosová kapacita systému s najnižšou degradáciou signálu využila čo najlepšie, dôležitým aspektom sa stáva dôkladné pochopenie, modelovanie a charakterizovanie správania viackanálových optických systémov v rôznych prevádzkových prostrediach.

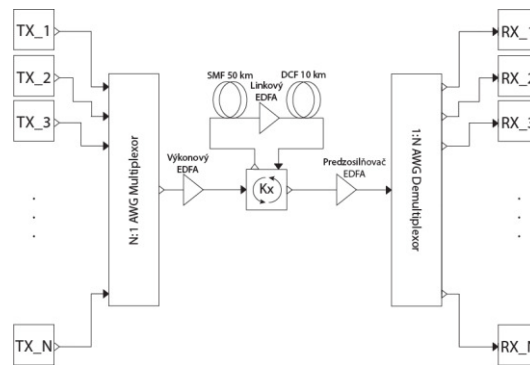
Je zrejme, že konečným cieľom je návrh úplne transparentných optických sietí, pretože v súčasnosti jedine tieto siete ponúkajú najekonomickejšie riešenie s transparentnosťou v bitovej rýchlosti, formáte signálu a protokolu. Transparentné optické siete však čelia mnohým výzvam, pričom hlavnými problémami sú problémy s kontrolou a riadením týchto sietí a potreba správy koncového inžinierstva spôsobeného nahromadením fyzických zariadení v optickej linke.

## II. OPTICKÉ VLÁKNOVÉ PRENOSOVÉ SYSTÉMY S HUSTÝM VLNOVÝM DELENÍM

Optické systémy s hustým vlnovým delením DWDM sú veľmi sľubným riešením pre viackanálové optické vlákňové komunikácie. Referenčná vlnová dĺžka pre DWDM kanály je 1552,5 nm (193,10 THz). Dnešné komerčne dostupné DWDM systémy dokážu multiplexovať až 128 optických nosných. Pri maximálnej prenosovej rýchlosti 2,5 Gb/s možno multiplexovať až 128 kanálov a až 40 kanálov s maximálnou prenosovou rýchlosťou 10 Gb/s. DWDM systém je schopný zvládnuť prenosy s prenosovými rýchlosťami 10 Gb/s, 20 Gb/s, ale teoreticky aj 40 Gb/s a 100 - 112 Gb/s.

Poloha kanálov DWDM je 1 624,8914 nm (184,5 THz) až 1 530,0413 nm (195,9375 THz) a medzikanálové rozostupy sú 0,1 nm (12,5 GHz), 0,2 nm (25 GHz), 0,4 nm (50 GHz) a 0,8 nm (100 GHz) pri normalizačnej frekvencii 193,1 THz. V súčasnej dobe sú v DWDM systémoch využívané medzikanálové rozostupy 100 GHz a 50 GHz v závislosti od šírky pásma kanálov. Nižšie rozostupy stále nie sú v chrbitcových optických sieťach veľmi rozšírené [1], [2].

Typický príklad optickej siete s DWDM je znázornený na Obr. 1. Prenosová trasa optického systému s DWDM je zvyčajne tvorená kaskádou identicky zapojených častí obsahujúcich vhodné bloky na úpravu optického signálu.



Obr. 1 Schéma DWDM systému.

Na strane prijímača sa nachádza laserový zdroj generujúci žiarenie, na ktoré sa moduluje elektrický signál do optického dátového toku. Niekoľko takýchto optických tokov sa následne multiplexuje a prenáša optickou trasou tvorenou jednovidovým optickým vláknom (SMF, Single Mode Fiber) a optickým zosilňovačom. Optický výkonový zosilňovač neguje vložné straty multiplexora na strane vysielača a optický predzosilňovač zosilňuje optický signál tak, aby bol detekovateľný prijímačom. Na dlhých prenosových trasách je zvyčajné aj zahrnutie optického linkového zosilňovača, ktorého hlavnou úlohou je kompenzovať straty spôsobené prenosom cez optickú trasu. Signál je demultiplexovaný a následne šírený do jednotlivých prijímačov, ktoré obnovujú pôvodné elektrické dáta detekciou a demoduláciou optického signálu [3].

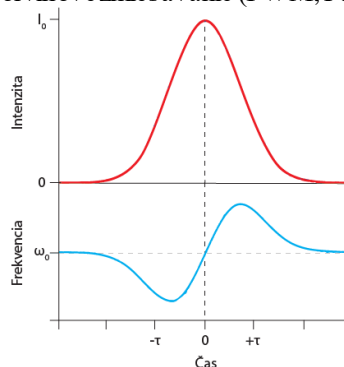
### III. NELINEÁRNE JAVY V DWDM SYSTÉMOCH

Lineárne javy sú hlavným obmedzením pre optické prenosové systémy operujúce s nízkym vysielačím výkonom, pri prenose na nízkej vzdialenosti a prenosovej rýchlosti 2,5 Gb/s na kanál. Avšak pri vyšších prenosových rýchlostiach (nad 10 Gb/s) narážame na obmedzenia spojené s nelineárnymi javmi.

Rozlišujeme dva typy nelineárných javov. Prvý typ nelineárných javov vzniká v dôsledku interakcie svetelných vln s fonónmi (molekulárnymi vibráciami) v prostredí oxidu kremičitého. Do tejto kategórie zaraďujeme nelineárne rozptyľové javy – stimulovaný Ramanov rozptyľ a stimulovaný Brillouinov rozptyľ. Druhým typom nelineárných javov sú nelineárne javy vznikajúce v dôsledku závislosti indexu lomu optického vlákna na intenzite svetelného lúča šíriaceho sa optickým vláknom [4], [5]. Túto závislosť tiež nazývame Kerrov efekt. Index lomu jadra optického vlákna je možné vyjadriť ako:

$$n = n_0 + n_2 \cdot \frac{P}{A_e}, \quad (1)$$

kde  $n_0$  je index lomu jadra optického vlákna,  $n_2$  je tzv. nelineárny index lomu,  $P$  je výkon naviazaný na optické vlákno a  $A_e$  je efektívna plocha jadra optického vlákna. Kerrov efekt sa prejavuje v rozličných formách. Na základe tvaru a formy vstupného signálu sa môže prejavovať ako vlastná fázová modulácia (SPM, Self Phase Modulation), krížová fázová modulácia (XPM, Cross Phase Modulation) a štvorvlnové zmiešavanie (FWM, Four Wave Mixing) [4].



Obr. 2 Vplyv SPM na prenášaný impulz.

#### A. Vlastná fázová modulácia

SPM je dôsledkom Kerrovho efektu a vzniká v dôsledku závislosti intenzity svetla naviazaného na optické vlákno od indexu lomu optického vlákna. Táto závislosť spôsobuje posun fázy signálu

v závislosti na jeho intenzite, a teda zmenu frekvenčného spektra (Obr. 2). V strede grafu má frekvencia signálu takmer lineárny charakter. Ako sa výkon signálu zvyšuje, index lomu optického vlákna sa tiež zvyšuje a frekvencia signálu vplyvom SPM klesá. Naopak, ak sa výkon signálu znižuje, znižuje sa aj index lomu optického vlákna a frekvencia signálu rastie. Tento jav sa nazýva frekvenčný chirping.

Zmena fázy optického signálu vplyvom SPM je teda závislá od intenzity prenášaného signálu. Efekt chirpingu je obzvlášť výrazný v optických systémoch využívajúcich vysoké vysielacie výkony. Keďže sa SPM vyskytuje v úzkom prepojení s chromatickou disperziou, je dôležitý jav SPM zohľadňovať v systémoch s vysokou prenosovou rýchlosťou, ktoré sú chromatickou disperziou obzvlášť limitované. V systémoch s prenosovou rýchlosťou nad 10 Gb/s a v systémoch s vysokým vyžarovacím výkonom SPM výrazne zvyšuje účinky chromatickej disperzie, teda prekrývanie prenášaných optických impulzov [6], [7], [8].

### B. Krížová fázová modulácia

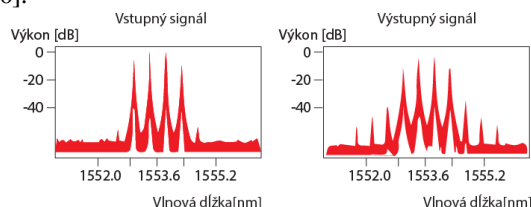
XPM je ďalšou formou Kerrovhovho efektu a principiálne vychádza z vlastnej fázovej modulácie. Kvôli eliminácii ostatných nelineárnych javov je výhodné analyzovať vplyv SPM v jednonábovom systéme. XPM sa prejavuje pri prenose dvoch alebo viacerých optických signálov alokovaných v tesnej blízkosti. Preto je XPM dôležitý zohľadňovať v DWDM systémoch. SPM a XPM sú si v ohľade závislosti indexu lomu optického vlákna od intenzity optického signálu veľmi podobné. Avšak v prípade XPM je celkový fázový posun prenášaného optického impulzu v jednom kanáli ovplyvnený vlastnosťami susedných kanálov. Index lomu optického vlákna je tiež určený celkovou intenzitou všetkých prenášaných kanálov. Krížová fázová modulácia v skutočnosti spôsobuje, že kolísanie výkonu v určitej vlnovej dĺžke sa prenesie do fázového kolísania v ostatných šíriaciach sa kanáloch. Výsledkom XPM môže byť nesúmerné rozšírenie spektrálnej čiary a deformácia tvaru impulzu [4], [8].

Vplyv XPM na viacnábový optický komunikačný systém sa prejavuje vo väčšej miere ako SPM a jeho degradačné účinky narastajú so zvyšujúcou sa intenzitou optického žiarenia naviazaného na optické vlákno. Rozhodujúcim faktorom je aj rozstup medzi jednotlivými kanálmi v systémoch DWDM. Čím je rozstup medzi kanálmi menší, tým väčší má XPM vplyv na kvalitu preneseného signálu [8].

### C. Štvorvlnové zmiešavanie

V DWDM systéme s kanálmi s uhlovými frekvenciami  $\omega_1, \omega_2, \dots, \omega_n$  závislosť intenzity prenášaného signálu od indexu lomu optického vlákna vyvoláva nielen fázové posuny v rámci kanálov, ale taktiež spôsobuje vznik signálov na nových frekvenciách, blízkych pôvodným kanálom, ako napr.  $2\omega_i - \omega_j$  a  $\omega_i + \omega_j - \omega_k$ . Tento jav sa nazýva štvorvlnové zmiešavanie. Na rozdiel od javov SPM a XPM, ktoré sú významné pre systémy s vysokou prenosovou rýchlosťou, vplyv FWM je nezávislý od prenosovej rýchlosti. Naopak, vplyv FWM je kriticky závislý od medzikanálových rozstupov DWDM systému a od chromatickej disperzie. Znižovanie medzikanálových rozstupov, ale aj znižovanie chromatickej disperzie, zvyšuje vplyv FWM [4], [9]. Účinky FWM teda musia byť brané v úvahu aj pre systémy s nízkymi prenosovými rýchlosťami. Vplyv FWM na spektrum prenášaného signálu je zobrazený na Obr. 3.

V DWDM systémoch sa často používajú optické vlákna s nenulovou disperziou, a to z dôvodu, že pre optické vlákna s nulovou disperziou je vplyv FWM kritický. FWM je degradačný jav, ktorý výraznou mierou obmedzuje realizovateľnosť DWDM systémov. FWM možno využiť v aplikáciách, pri ktorých je FWM využité pri meraní nelinearity a chromatickej disperzie v optických vláknach [10].

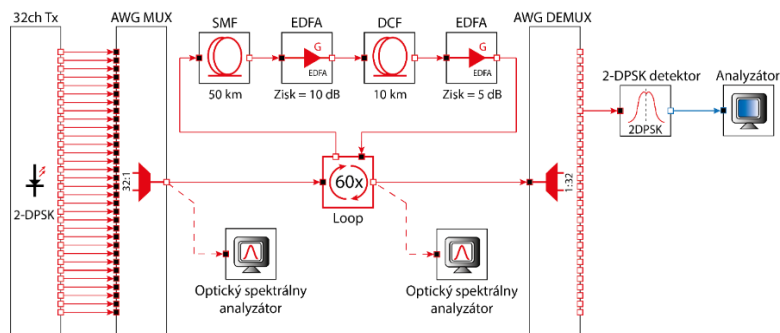


Obr. 3 Vplyv FWM na prenášaný impulz.

## IV. NÁVRH ULTRA-DWDM SYSTÉMU S VPLYVOM NELINEÁRNYCH JAVOV SPM, XPM A FWM

Na základe nášho dlhodobého výskumu a naakumulovaných poznatkov a vedomostí sme dospeli k návrhu kompletného Ultra-DWDM systému s kapacitou 0,32 Tb/s (10 Gb/s na 1 kanál), 1,28 Tb/s (40 Gb/s na 1 kanál) a 3,2 Tb/s (100 Gb/s na 1 kanál). Celková dĺžka prenosovej trasy tvorenej jednovidovým optickým vláknom je 3 600 km (60×(50 km SMF + 10 km DCF)). Schéma

navrhnutého systému je na Obr. 4. K zosilneniu signálu dochádza po každom 50. kilometri o hodnotu 10 dB a 60. kilometri o hodnotu 5 dB. Zosilňovač je typu EDFA s obojsmerným režimom čerpania na vlnovej dĺžke 1480 nm. Prenosová rýchlosť kanála je 40 Gb/s a 100 Gb/s, typ použitej modulácie je 2-DPSK. Ďalšie parametre tohto modelu sú uvedené v Tab. 1.

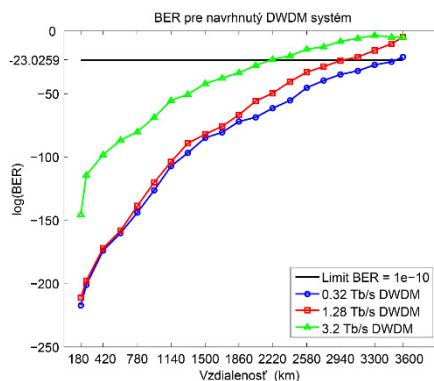


Obr. 4 Navrhnutý model Ultra-DWDM systému.

Tabuľka 1  
Parametre navrhnutého Ultra-DWDM modelu

Vysielač - 32ch Tx	
Bitová rýchlosť	10 Gb/s
	40 Gb/s
	100 Gb/s
Počet vzoriek na 1 bit	64
Počiatočná frekvencia	193,0 THz
Medzikanálové rozostupy	100 GHz
Vyžarovací výkon	10 dBm
Typ modulácie	2-DPSK
AWG multiplexor	
Počet portov	32
Útlm na 1 port	5,5 dB
Prenosová trasa	
Dĺžka SMF	50 km
Hodnota chromatickej disperzie SMF	17 ps/nm/km
Referenčná vlnová dĺžka	1550 nm
Útlm	0,2 dB/km
Disperzia - $\beta_2$	-20 ps <sup>2</sup> /km
Polarizačná disperzia - PMD	0,2 ps/km
Efektívna oblasť vlákna $A_e$	80 $\mu\text{m}^2$
Nelineárny index lomu $n_2$	2,6e-020 m <sup>2</sup> /W
Dĺžka DCF	10 km
Hodnota chromatickej disperzie DCF	-85 ps/nm/km
Zisk EDFA 1	10 dB
Zisk EDFA 2	5 dB
Celková dĺžka prenosovej trasy	3600 km
AWG demultiplexor	
Počet portov	32
Útlm na 1 port	5,5 dB

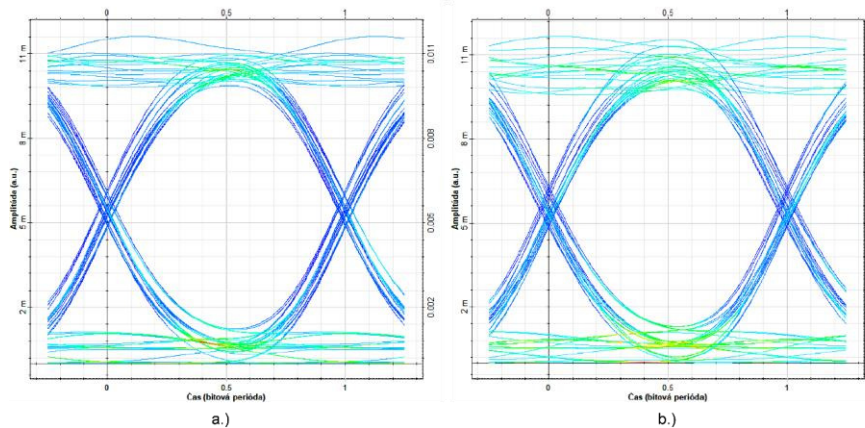
Po vykonaní simulácií pri rôznych prenosových rýchlostiach sme zaznamenali dosiahnuté hodnoty BER do grafickej závislosti na Obr. 5. Súčasný limit BER pre DWDM systémy je  $1e-10$ , čo po zlogaritmovaní predstavuje hodnotu -23,0259. To znamená, že DWDM systém je z hľadiska spoľahlivosti nevyhovujúci ak je hodnota BER prijatého signálu vyššia ako  $1e-10$  (resp. -23,0259).



Obr. 5 Maximálny dosah DWDM systému s rôznou prenosovou kapacitou.

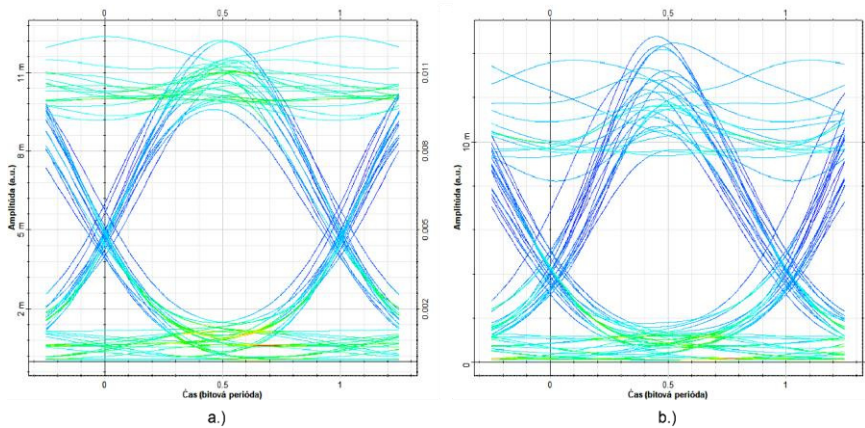
Maximálny dosah navrhnutého DWDM systému s kapacitou 0,32 Tb/s (10 Gb/s na kanál) je na základe dosiahnutého BER =  $1,69 \times 10^{-11}$  (Q-faktor = 6,6290 dB) 3 480 km. Zvýšením kapacity na 1,28 Tb/s (40 Gb/s na kanál) sa maximálny dosah znížil na 2 940 km, čo zodpovedá hodnote BER =  $6,32 \times 10^{-11}$  (Q-faktor = 6,4314 dB). Pri prenosovej rýchlosti 100 Gb/s na kanál, teda celkovej kapacite DWDM systému 3,2 Tb/s, sa dosah podstatne znížil, a to na 2 220 km. To zodpovedá hodnote BER =  $1,37 \times 10^{-10}$  (Q-faktor = 6,3128 dB).

Nasledujúce diagramy oka, na Obr. 6, korešpondujú s 1,28 Tb/s DWDM systémom. Z daných diagramov sme vypočítali hodnoty BER a Q-faktoru: BER =  $1,27 \times 10^{-33}$ , Q-faktor = 12,0277 dB pre 1 680 km a BER =  $2,92 \times 10^{-22}$ , Q-faktor = 9,6323 dB pre 2220 km.



Obr. 6 Diagram oka po trase a.) 1680 km, b.) 2220 km pre 1,28 Tb/s DWDM.

Obr. 7 korešponduje s 3,2 Tb/s DWDM systémom. Môžeme si všimnúť, že s narastajúcou vzdialenosťou sa otvorenosť oka znižuje a skreslenie sa zvyšuje. Z daných diagramov sme vypočítali hodnoty BER a Q-faktoru: BER =  $9,39 \times 10^{-25}$ , Q-faktor = 10,2052 dB pre 1 140 km a BER =  $1,46 \times 10^{-12}$ , Q-faktor = 6,9815 dB pre 2 040 km.



Obr. 7 Diagram oka po trase a.) 1140 km, b.) 2040 km pre 3,2 Tb/s DWDM.

## V. ZÁVER

Na základe dosiahnutých výsledkov možno konštatovať, že zmenou tvaru vstupného signálu je možné vo veľkej miere ovplyvniť vznik a vplyv disperzie skupinovej rýchlosti označovanej ako GVD (Group Velocity Dispersion) a nelineárnych javov v DWDM systémoch. Výskum v tejto oblasti je nesmierne dôležitý a umožňuje dosahovať stále vyššie prenosové kapacity a prenosové vzdialenosti DWDM systémov. Existujúce optické prenosové systémy využívajúce OOK (On-Off Keying) modulácie môžu byť pomerne jednoducho modifikované použitím duobinárnej modulácie DRZ (Duobinary Return to Zero) alebo modifikovanej duobinárnej modulácie MDRZ (Modified Duobinary Return to Zero), bez potreby zmeny prijímacích modulov. Takýmto spôsobom je možné prekonať väčšie vzdialenosti vďaka vyššej odolnosti voči XPM a FWM.

## POĎAKOVANIE

Táto práca bola podporená výskumnými projektami VEGA 1/0584/20 “UWB senzorové systémy monitorovania osôb pracujúce v reálnych podmienkach (ReConLoc)” a FEI-2021-71 “Metódy zabezpečenia dát v moderných komunikačných systémoch”.

## LITERATÚRA

- [1] N. Antoniadis, et al., “*WDM Systems and Networks, Modeling, Simulation, Design and Engineering*”, 2012, 549 pp.
- [2] ITU-T G.694.1. Series G: Transmission Systems and Media, Digital Systems and Networks: Transmission media and optical systems characteristics – Characteristics of optical systems. Spectral grids for WDM applications: DWDM frequency grid. 2012.
- [3] A. Dochhan, et al., “Solutions for 80 km DWDM Systems”, *Journal of Lightwave Technology*, 2015, pp. 491–499.
- [4] G. P. Agrawal, “*Nonlinear Fiber Optics*”, 2019, 631 pp.
- [5] S. P. Singh, et al., “Nonlinear Effects in Optical Fibers: Origin, Management and Applications”, *Progress in Electromagnetic Research*, 2007, pp. 249-275.
- [6] R. Ramaswami, et al., “*Optical Networks*”, 2010, 928 pp.
- [7] T. Ivaniga, et al., “Investigation of SPM in WDM System with EDFA”, *IEEE Carpathian Journal of Electronic and Computer Engineering*, 2016, pp. 7-12.
- [8] T. Ivaniga, et al., “Ovplyvňovanie javu CPM s využitím EDFA v CWDM systéme”, *Electrical Engineering and Informatics*, 2017, pp. 62-66.
- [9] T. Ivaniga, et al., “The Effect of Four-Wave Mixing to the Four-Channel DWDM System with Spacing According to the ITU-T G.694.1”, in *Informatics*, 2015, pp. 128-132.
- [10] T. Ivaniga, et al., “The Influence of FWM with AWG Multiplexor in DWDM System”, *Przegląd Elektrotechniczny*, 2018, pp. 113-117.



# Rozšírenie Minecraftu z pohľadu fungovania Bitcoinu

<sup>1</sup>Daniel Florek <sup>2</sup>Anton Baláž <sup>3</sup>Branislav Madoš

Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, Katedra počítačov a informatiky

<sup>1</sup>daniel.florek@student.tuke.sk <sup>2</sup>anton.balaz@tuke.sk <sup>3</sup>branislav.mados@tuke.sk

**Abstrakt – Témou tohto článku je implementácia modifikácie do hry Minecraft, ktorá by jednoduchou a zábavnou formou demonštrovala základné používateľské koncepty kryptomeny Bitcoin.**

## I. ÚVOD

Bitcoin je fenoménom posledných rokov. Hoci vznikol už v roku 2008 popularitu začal naberať až na prelome rokov 2013 a 2014, kedy jeho cena prvý krát presiahla hodnotu 100 dolárov. Od tej doby cena Bitcoinu neustále rastie a v čase písania tohoto dokumentu stála na hodnote 60500 dolárov za jeden Bitcoin.

Bitcoin má mnoho podporovateľov a mnoho odporcov. Veľa ľudí sa však oň nezaujíma zo strachu, nedostatku informácií, či potrebe investovať skutočné peniaze. Cieľom tohto článku je práve táto skupina ľudí, pretože výstup práce dokáže jednoduchým, vo forme hry, zábavným spôsobom predstaviť fungovanie siete Bitcoin z používateľského hľadiska bez rizika straty peňazí.

## II. PEER-TO-PEER

Pri študovaní Bitcoinu je najskôr potrebné sa pozrieť na Peer-to-Peer komunikáciu, pretože je nad ňou postavená komunikácia a koncepty celej siete Bitcoin.

### A. Topológia

Peer-to-Peer (P2P) [1] je alternatívna topológia k tradičnej klient-server topológii. Miesto získavania obsahu z centrálného serveru si obsah navzájom distribuujú všetci účastníci siete. Zariadenia v sieti sa nazývajú **uzly** a zastávajú rolu aj servera aj klienta. Uzly si medzi sebou vymieňajú nielen obsah, ale aj "kontakt" (IP adresy) na ďalšie uzly, ktoré poznajú.

Výhoda Peer-to-Peer topológie je v tom, že odstraňuje nápor na centrálny server, pretože si uzly obsah distribuujú medzi sebou, takže je to rýchlejšie a efektívnejšie. Na druhú stranu však zvyšuje nápor na sieť, pretože Peer-to-Peer služby častokrát vytvárajú stovky až tisíce obojsmerných pripojení. Navyše tieto služby častokrát využívajú protokol UDP, ktorý nevytvára stále pripojenie, takže v rámci réžie komunikácie medzi uzlami sa odošle veľké množstvo pomocných paketov.

Väčšina týchto sietí však využíva pomocné servery, ktoré sa využívajú pre nadväzovanie komunikácie medzi uzlami. Keď sa pripája nový uzol do siete, tak takýto server kontaktuje a ten mu poskytne počiatočný zoznam IP adries uzlov, ktoré server pozná. Tieto uzly už následne uzol kontaktuje sám a začne sa sám zapájať do fungovania služby.

Peer-to-Peer si zo začiatku osvojili hlavne systémy pre zdieľanie súborov, ako napríklad BitTorrent, no časom si ich vo veľkom osvojili aj siete kryptomien, pretože hlavným prvkom väčšiny kryptomien je práve decentralizovanosť, aby menu nemohol nikto regulovať alebo s ňou inak manipulovať.

### B. Peer-to-Peer koncepty v Bitcoine

Zariadenia v sieti Bitcoin sa aj v tomto prípade volajú **uzly** [2], [3]. Všetky uzly v sieti sú si rovné, to znamená, že neexistujú nadradené a podradené uzly, avšak nie každý uzol vykonáva rovnakú úlohu. Role, ktoré môžu jednotlivé uzly v sieti Bitcoin zastávať sú:

- **Network Routing Node** - jedná sa o takmer všetky uzly. Úlohou týchto uzlov je poskytovať IP adresy známych uzlov na vyžiadanie. A distribúcia dát, ktoré uzol obdrží (napríklad transakcie) ďalej do siete.

- **Full Blockchain Node** - uzly, ktoré u seba uchovávajú celú kópiu Blockchainu. Spravujú ho a distribuujú ďalším uzlom na vyžiadanie.
- **Miner Node** - uzly, ktoré vykonávajú proces ťaženia. Využívajú špeciálny software určený pre hromadnú kontrolu transakcií a súťažia o to, kto získa Bitcoin. Tieto uzly sú častokrát aj *Full Blockchain* uzly.
- **Wallet Node** - uzly reprezentujúce koncových používateľov. Tieto uzly vytvárajú transakcie, ktoré zvyšok siete kontroluje, organizuje do blokov a ukladá do Blockchainu.

K tomu, aby sieť dokázala správne fungovať, tak potrebuje mať spôsoby, ktorými sa uzly dokážu navzájom nájsť, kontaktovať a vymieňať si dôležité dáta. K tomu slúži niekoľko počiatočných spôsobov hľadania uzlov:

- **DNS seedy** - sú to servery, ktoré indexujú IP adresy uzlov. Častokrát na týchto serveroch bežia crawlery Bitcoin siete k tomu, aby zistili IP adresy uzlov
- **Seed adresy** - uzly u ktorých sa predpokladá, že budú vždy online

Po tom, čo uzol má určitú počiatočnú množinu ďalších uzlov, tak sa využívajú ďalšie spôsoby distribúcie IP adres.

- **Požiadavka getaddr** - jeden uzol si od druhého uzlu vyžiada adresy ďalších uzlov. Žiadaný uzol vyberie maximálne 2500 adres a vráti ich žiadajúcemu uzlu
- **Odoslanie svojej IP adresy** - raz za 24 hodín uzol odošle svoju IP adresu známym uzlom, ktoré ju rozdistribuujú ďalej
- **Nepodmienená distribúcia IP adres** - uzol môže náhodne od iného uzla obdržať zoznam adres. Uzol si takéto IP adresy uloží a zo zoznamu známych adres náhodne vyberie vyhovujúce adresy a odošle ich dvom ďalším náhodne zvoleným uzlom

Uzol si uchováva IP adresy známych uzlov v lokálnej databáze. Túto databázu upravuje keď si vyžiada zoznam IP adres uzlov, keď sa nevie s nejakým uzlom kontaktovať, alebo keď obdrží zoznam adres uzlov nepodmienené.

### III. BITCOIN

Táto sekcia sa v krátkosti zaoberá základnou funkcionalitou Bitcoinu. V prípade záujmu o detailnejší popis Bitcoinu, zdroj [4].

#### A. Pár kryptografických kľúčov a Bitcoin adresa

Základom celej funkcionality Bitcoinu je pár kryptografických kľúčov [2], [3], [4]. To znamená privátny a verejný kľúč.

Privátny kľúč je náhodné 256 bitové číslo. Väčšinou sa generuje vložení náhodného množstva bitov do SHA-256 hashovacej funkcie, čím sa vyprodukuje 256 bitový privátny kľúč.

Verejný kľúč sa produkuje z privátneho kľúča pomocou eliptickej krivky definovanej v štandarde **Secp256k1**. Privátny kľúč sa vynásobí s preddefinovaným generátorovým bodom **G**. Výsledkom takejto akcie je bod na krivke **K**, ktorý predstavuje verejný kľúč. Hodnoty **X** a **Y** takéhoto bodu sú 256 bitové čísla.

Z Verejného kľúča sa potom pomocou SHA-256 a RIPEMD-160 hashovaní vytvorí Bitcoin Adresa, ktorá sa zakóduje pomocou Base58Check kódovania. Táto Bitcoin Adresa slúži pre identifikáciu prijímateľa v transakciách (podobne ako číslo účtu).

#### B. Bitcoin Wallet

Bitcoin Wallet (Bitcoin peňaženka) sú programy, ktoré v základnej funkcionalite dokážu všetky spravovať kryptografické kľúče, vytvárať transakcie a zistiť zostatok Bitcoinov. Ďalšia funkcionálna sa líši v závislosti od konkrétnej implementácie [2], [3], [4].

Podľa platformy na ktorej fungujú je možné Bitcoin peňaženky rozdeliť na:

- **Desktop peňaženky** - prvé peňaženky vznikali práve na desktopy.
- **Mobilné peňaženky** - dostupné, vhodné pre začiatočníkov
- **Hardwarové peňaženky** - hardwarové zariadenia, na ktoré sa ukladajú kľúče a s počítačom/mobilom komunikujú cez USB alebo NFC
- **Webové peňaženky** - taktiež pomerne dostupné a vhodné pre začiatočníkov, avšak najnebezpečnejšie, pretože užívateľ nemá pod kontrolou svoje kľúče

Podľa autonómnosti voči sieti Bitcoin sa peňaženky rozdeľujú na:

- **Full-node** - má stiahnutú celú kópiu Blockchainu u seba, takže dokáže všetky transakčné úkony vykonávať sama
- **Lightweight** - nemá stiahnutú kópiu Blockchainu. Ukladá si iba pár transakcií a kontaktuje Full-node uzly k tomu, aby získala informácie o transakciách. Sieť Bitcoin však kontaktuje priamo

- **Third-party API** - nemá stiahnutú kópiu Blockchainu a informácie zo siete Bitcoin čerpá cez API rozhrania

Podľa spôsobu spravovania kľúčov sa peňaženky rozdeľujú na:

- **Nedeterministické** - jednotlivé prívátne kľúče sa generujú náhodne a každý kľúč sa musí spravovať a zálohovať samostatne
- **Deterministické** - kľúče sú vygenerované zo seedu, takže sa dajú ľahko zálohovať, avšak pri ukradnutí seedu sa získa prístup ku všetkým kľúčom ktoré sú z neho vygenerované
- **Hierarchické Deterministické (HD)** - kľúče sa organizujú do stromovej štruktúry, kedy z každého kľúča je možné vygenerovať niekoľko ďalších kľúčov, z ktorých je možné vygenerovať ďalšie kľúče až do neobmedzenej úrovne zanorenia. Koreňový kľúč stromu sa vygeneruje pomocou seedu

### C. Transakcie

Sieť Bitcoin k uchovávaniu stavu Bitcoinov na jednotlivých Bitcoin Adresách využíva systém transakcií uložených v Blockchaine [2], [3], [4]. Každá transakcia sa skladá zo vstupov a výstupov. Výstupy transakcie definujú množstvo posielaných Bitcoinov a takzvaný zamykací skript v jazyku **Script**, ktorý definuje ako sa k týmto Bitcoinom dá prístup. Väčšina zamykacích skriptov výstupy zamykajú práve na Bitcoin Adresu a digitálny podpis, kedy užívateľ, ktorý chce Bitcoin minúť musí dodať verejný kľúč, ktorý sa zahashuje do Bitcoin Adresy vo výstupe a digitálny podpis, ktorý je možné takýmto verejným kľúčom overiť. Výstupy, ktoré ešte nie sú minuté sa nazývajú **unspent transaction output (UTXO)**.

Vstupy novej transakcie referencujú výstupy existujúcich transakcií, ktoré v čase vytvárania ešte neboli minuté. Pri každom vstupe sa uvádza takzvaný odomykací skript, ktorý slúži ako vstup pre zamykací skript v referencovanom výstupe.

Zostatok Bitcoinov na určitej Bitcoin Adrese predstavuje súčet všetkých neminutých výstupov, ku ktorým má daná Bitcoin Adresa prístup.

### D. Verifikácia transakcií

K tomu, aby mohla byť nová transakcia zaradená do Blockchainu, sa musí overiť, či tvorca transakcie je skutočne oprávnený minúť Bitcoin z výstupov, ktoré referencuje na vstupoch novej transakcie [2], [3], [4]. Verifikácia sa vykoná interpretovaním skriptu, ktorý sa vytvorí spojením odomykacieho skriptu zo vstupu novej transakcie a zamykacieho skriptu z referencovaného výstupu existujúcej transakcie. V prípade, že interpretácia skriptu pre každý transakčný vstup skončí v poriadku, tak je takúto transakciu možné zaradiť do Blockchainu.

Transakčné skripty sú pomerne rozsiahla téma a je možné sa o nich dočítať viac v jednotlivých literárnych odkazoch.

### E. Blockchain

Blockchain [2], [3], [4] je štruktúra, ktorú sieť Bitcoin využíva k ukladaniu transakcií. Majú ho u seba uložené Full-Blockchain uzly a tieto uzly ho na vyžiadanie distribuujú ďalším uzlom.

Blockchain sa skladá z **Blokov**. Blok sa ďalej skladá z hlavičky, ktorá obsahuje metadáta a z listu platných transakcií. Každý blok v hlavičke nesie ukazateľ na predchádzajúci blok, ktorý bol v čase jeho vzniku posledný. V súčasnosti má celý Blockchain veľkosť okolo 330GB. Z celého Bitcoinu sa práve Blockchain považuje za najrevolučnejšiu technológiu.

### F. Ťažba (mining) a Proof-of-Work

Ťažba [2], [3], [4] predstavuje proces verifikácie transakcií a správy Blockchainu. Do siete sú pripojené uzly, ktoré majú spustený špeciálny ťažiaci software. Tieto uzly prijímajú transakcie zo siete, kontrolujú ich a platné transakcie si ukladajú do svojho dočasného lokálneho bloku.

Proof-of-Work [2], [3], [4] je algoritmus pomocou ktorého sa vyberá uzol, ktorý získa odmenu za ťaženie v podobe nových Bitcoinov. V rámci siete Bitcoin je definovaná globálna 256 bitová číselná hodnota **Difficulty Target**. Ťažiacie bloky vždy, keď vykonajú nejakú zmenu v rámci svojho lokálneho bloku, tak vytvoria SHA-256 hash z jeho hlavičky. Tento hash porovnávajú s hodnotou Difficulty Targetu a ak je hash menší alebo rovný ako Difficulty Target, tak je možné ho považovať za vyťaženie. Uzol takýto blok potom rozošle zvyšku siete, kde si ho každý uzol skontroluje a ak je v poriadku, tak si ho uloží do svojej kópie Blockchainu a pošle ďalej do siete. Týmto spôsobom dokáže Bitcoin fungovať decentralizovaným spôsobom a stále obsahovať iba platné transakcie.

Prvá transakcia v bloku sa nazýva **Coinbase** transakcia, ktorá obsahuje odmenu pre uzol, ktorý blok vyťažil. Coinbase sa skladá z nových Bitcoinov vypustených do obehu a z poplatku za prevod zo všetkých transakcií v danom bloku.

#### IV. MINECRAFT

Minecraft je hra, ktorú v roku 2009 vydal Švéd Markus Persson vystupujúci pod pseudonymom "Notch". Na hre spočiatku pracoval sám a neskôr v roku 2011 založil štúdio Mojang. Hru spolu so štúdiom v roku 2014 predal Microsoftu. Pôvodná hra bola naprogramovaná v jazyku Java. Po odkúpení hry Microsoftom vznikla aj alternatívna verzia s názvom Minecraft Windows 10 Edition, ktorá je preprogramovaná do jazyku C++ a je exkluzívna pre platformu Windows 10. Avšak Minecraft Java Edition je jediná verzia, ktorá podporuje tvorbu komunitných modifikácií.

Minecraft je jednoduchá a zároveň komplexná hra. Odohráva sa v náhodne procedurálne vygenerovanom svete v ktorom je všetko z kociek a jej jedinou náplňou je prežiť. K tomu ponúka veľké množstvo mechaník, od záhradkárčenia, chovu zvierat, ťažby materiálov, výroby zbraní a nástrojov až po koncepty elektrických obvodov a je len na hráčovi samotnom, ako sa v tomto virtuálnom svete zabaví.

##### A. Modifikácia Minecraftu

K modifikácií Minecraftu sa využívajú externé modifikačné API, ktoré umožnia prepojiť modifikácie vytvorené komunitou so samotnou hrou. Príkladom takéhoto API je aj Forge. Ten treba stiahnuť z oficiálnej stránky a nainštalovať do priečinku herného klienta.

Forge [5] v rámci hry poskytuje dve zbernice udalostí: **Módová zbernica** a **Forge zbernica**. Keď sa hra zapína, tak Forge prechádza cez všetky Java Archívy (.jar) v priečinku **mods** v priečinku s hernými súbormi. V týchto Java Archívoch hľadá triedy označené anotáciou **@EventBusSubscriber**, ktorá určuje, na ktorú zbernicu tieto triedy načúvajú. V rámci týchto tried potom hľadá metódy označené anotáciou **@SubscribeEvent**, ktoré označujú, že tieto metódy sa majú použiť ako obsluha udalostí. K akej udalosti sa daná metóda vzťahuje určujú jej argumenty.

Módová zbernica poskytuje udalosti, ktoré nastávajú pri načítaní hry. Počas týchto udalostí modifikácie môžu zaregistrovať do hry nové herné bloky, herné objekty, GUI obrazovky a tak podobne. Forge zbernica poskytuje udalosti, ktoré nastávajú pri samotnej hre ako napríklad pripojenie hráča na server, zničenie bloku, úmrtie hráča atď.

Modifikácia sa vytvára ako Gradle projekt, ktorý je možné stiahnuť taktiež z oficiálnej stránky Forge. Tento Gradle projekt stiahne knižnice Forge, ktoré implementujú špeciálne Java anotácie, ktorými je možné jednotlivé obsluhy registrovať na Forge zbernice pri spúšťaní hry.

Pri implementácii modifikácie sa pracuje so vstavanými triedami Minecraftu, ktoré boli získané dekompiláciou a deobfuskáciou hry.

##### B. Architektúra Minecraftu

Minecraft vždy beží na architektúre klient-server. Nezáleží na tom, či hráč hrá v móde jedného hráča (singleplayer) alebo v móde viacerých hráčov (multiplayer), vždy beží server, ktorý sa stará o správu herného sveta. Jediný rozdiel je v tom, že v prípade hry jedného hráča server beží na pozadí a nemajú k nemu prístup iní hráči a v prípade módu viacerých hráčov sa o svet stará dedikovaný server, ktorý je prístupný cez internet.

Klientom je herné okno, v ktorom hráč hrá hru. Server sa stará o herný svet - uchováva dáta o blokoch, spravuje inventáre/kontajnery, rieši pozíciu hráčov atď. Klient obsahuje len také dáta, ktoré sú potrebné pre render scény alebo pre zobrazenie GUI okien. Tieto dáta neustále čerpá zo serveru. Keď sa hráč prechádza po serveri, tak server klientovi odosiela informácie o blokoch, aké sa majú hráčovi vykresliť. Keď chce hráč otvoriť GUI okno bloku, tak klient odošle informáciu o použití bloku serveru a server mu cez internetový paket otvorí okno daného bloku a taktiež mu odošle obsah takéhoto bloku, ktorý sa má v okne zobraziť.

##### C. Limitácie Minecraftu

Minecraft prináša aj niekoľko obmedzení. Hlavným obmedzením je, že Minecraft renderuje grafické okná v malom rozlíšení, ktoré sa potom škáluje na rozlíšenie okna hry. Týmto spôsobom sa zabezpečuje jednoduchá pixelovitá grafika hry, no značne to obmedzuje priestor, na ktorom hra dokáže v rámci okien zobrazovať informácie.

Druhým obmedzením sú slabé možnosti samotných grafických rozhraní. Hra obsahuje malé množstvo vstupno/výstupných prvkov a aj tie, ktoré tam sú, tak sa s nimi pracuje ťažkopádne, pretože väčšina z nich bola vytváraná iba pre hlavné menu a nastavenia hry. Herné okná Minecraftu totiž nevyužívajú žiadne tlačidlá, textboxy, listboxy ani nič podobné. Preto, keď chce človek použiť tieto prvky v hernom okne, tak musí "ohýbať technológiu".

Ďalším obmedzením je obfuskácia kódu, kedy niektoré časti kódu majú nezmyselné názvy premenných, argumentov a metód. Komunita už vytvorila mapovania časti kódu na zmysluplné

názvy, avšak stále ostáva určitá časť kódu obfuskovaná, takže programátor musí zisťovať náplň funkcií prechádzaním riadok po riadku.

Posledným nedostatkom je nedostatočná dokumentácia modifikačného procesu. Minecraft ako proprietárny produkt neposkytuje žiadnu dokumentáciu a dokumentácia, ktorú poskytuje Forge API je veľmi slabá, pretože je stručná, neaktuálna a nevenuje sa názorným ukážkam kódu, takže pre začínajúceho programátora v oblasti modifikácie Minecraftu je prakticky nepoužiteľná.

## V. VYTVORENÉ RIEŠENIE

Cieľom praktickej časti práce bolo vytvoriť jednoduchú simuláciu Bitcoinu v Minecrafte, ktorá by slúžila pre vzdelávacie účely. Modifikácia do hry pridáva tri herné bloky: **Bitcoin Wallet Block**, **Mining Block** a **Blockchain Explorer Block**. Každý z týchto blokov poskytuje iné grafické rozhranie a simuluje iný aspekt siete Bitcoin. Vzhľad týchto blokov v hre je možné vidieť na obrázku 1.



Obr. 1 Vzhľad blokov v hre

### A. Transactions Capability (implementácia "Blockchainu")

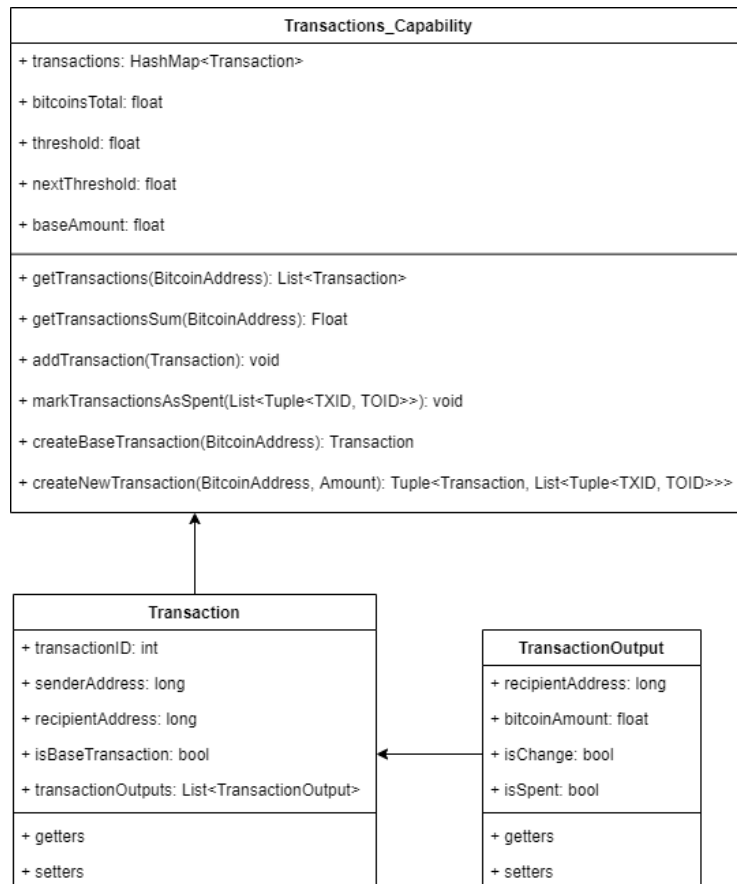
Trieda TransactionsCapability je jadrom celej modifikácie a implementuje veľmi zjednodušenú funkcionálnosť Blockchainu. Capabilities v rámci Minecraftu predstavujú spôsob, ktorým sa dá k existujúcej hernej entite pripojiť trieda, ktorá túto hernú entitu rozširuje o dodatočný stav bez potreby akokoľvek upravovať triedu hernej entity. Capabilities sú výhodné práve vtedy, keď je potrebné rozšíriť chovanie vstavaných herných entít Minecraftu. Inštancie Capabilities si vytvára aj klient aj server, avšak **dáta o capabilities si ukladá iba server**.

TransactionsCapability sa viaže k hernej entite (triede) **World**, takže k inštancii herného sveta, ktorý momentálne beží na serveri/klientovi.

Trieda TransactionsCapability uchováva transakcie v asociatívnom poli a implementuje metódy na vytváranie užívateľských transakcií, vytváranie Coinbase transakcií, či zisťovanie zostatku. TransactionsCapability **spravuje iba server**, t.j. vytvára nové transakcie. Keď hráči chcú vytvoriť transakciu, tak odošlú na server požiadavku, server požiadavku skontroluje a v prípade, že hráč žiada vytvoriť platnú transakciu, tak si ju server vytvorí u seba a rozošle o nej informáciu klientským aplikáciám hráčov. Klientské aplikácie si do svojej kópie TransactionsCapability takúto novú transakciu uložia, čím si udržiavajú presnú kópiu TransactionsCapability, ktorú má server.

TransactionsCapability na strane servera slúži pre perzistentné uchovávanie všetkých transakcií a pre kontrolu a vytváranie nových transakcií. TransactionsCapability na strane klienta slúži pre účely GUI, napríklad pre zobrazenie zostatku, či pre zobrazenie listu všetkých transakcií. Server aj klientské aplikácie všetkých hráčov majú vždy totožný stav svojich TransactionsCapabilities.

Diagram tried zobrazujúci implementáciu TransactionsCapability a kľúčových tried je možné vidieť na obrázku 2.



Obr. 2 Diagram tried TransactionsCapability

### B. Bitcoin Wallet Block

Bitcoin Wallet Block simuluje funkcionálnosť Bitcoin peňaženiek. Umožňuje generovať predmet typu **Bitcoin Key Pair**, ktorý v sebe nesie privátny kľúč a Bitcoin Adresu. Privátny kľúč sa vygeneruje ako náhodné číslo z rozsahu INT\_MAX po INT\_MAX \* 2. Bitcoin Adresa sa z privátneho generuje jednoduchou matematickou funkciou. Rozsah čísel sa oproti pôvodnej implementácii zvolil menší z toho dôvodu, že jediný spôsob ako sa dá v Minecrafte zobraziť informácia uložená v predmete je cez Tooltip. Keby sa využíval pôvodný 256 bitový rozsah, tak by sa takéto kľúče nemohli hráčovi zobraziť, pretože by vychádzali mimo obrazovku.

Každý hráč si môže vygenerovať ľubovoľný počet predmetov **Bitcoin Key Pair**, čím si vygeneruje viac Bitcoin Adries, na ktoré môže ťažiť, alebo prijímať Bitcoinov od iných hráčov.

Ďalej blok poskytuje rozhranie pre tvorbu transakcií. Obsahuje slot pre vloženie predmetu typu Bitcoin Key Pair. Po jeho vložení sa zobrazí zostatok Bitcoinov na Bitcoin Adrese vo vloženom predmete. Pod týmto slotom sa nachádzajú dve textové polia a tlačítko Send. Prvé textové pole umožňuje zadať množstvo posielať Bitcoinov. Druhé pole umožňuje zadať Bitcoin Adresu prijímateľa. Po kliknutí tlačítka Send klient odošle serveru žiadosť o vytvorenie novej transakcie. Ten skontroluje, či je takúto transakciu možné vytvoriť a v prípade, že áno, tak ju vytvorí a rozošle všetkým hráčom, ktorí si ju uložia do svojej kópie Transactions Capability.

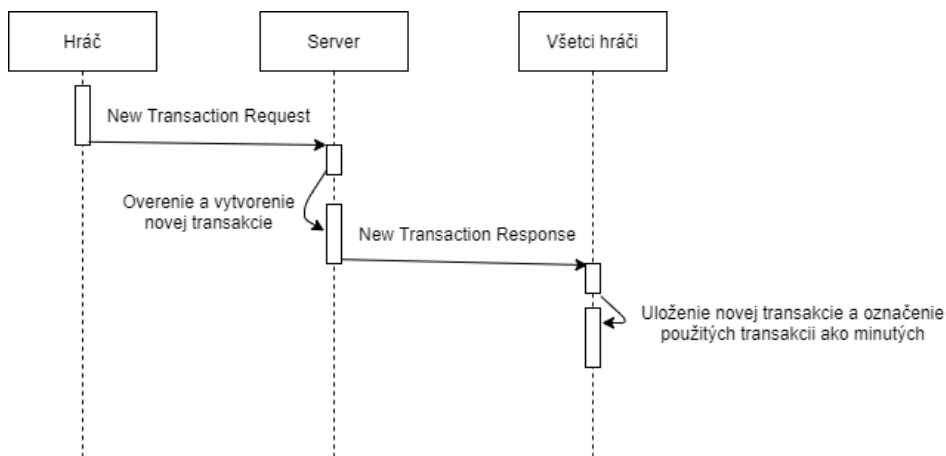
### C. Mining Block

Mining Block simuluje funkcionálnosť ťažobného softwaru. Vzhľadom k tomu, že transakcie kontroluje iba server a nevykonáva sa hromadná kontrola celou sieťou ako v prípade skutočného Bitcoinu, tak bolo potrebné vytvoriť alternatívu, ktorá by stále hráčom umožňovala získavať Bitcoinov.

Do grafického rozhrania Mining Blocku môže hráč vložiť predmet Bitcoin Key Pair a server periodicky každú minútu vyberie jeden Mining Block na ktorého Bitcoin Adresu vygeneruje Coinbase transakciu. Každý hráč si môže vytvoriť ľubovoľný počet Mining Blockov, čím sa môže zapojiť do ťažby viacnásobne a zvýšiť tak šancu, že niektorá z jeho Bitcoin Adries bude vybratá. Do Mining Blocku je taktiež možné vložiť diamanty (sú vstavená surovina Minecraftu, ktorá sa náhodne generuje v rámci sveta a hráči ju môžu ťažiť), ktoré danému Mining Blocku



Obr. 3 GUI Bitcoin Wallet Blocku



Obr. 4 Sekvenčný diagram vytvárania novej transakcie

zvyšujú šancu na získanie Bitcoinov. Maximálny počet diamantov, koľko sa do jedného bloku dá vložiť je 64. Koľko diamantov sa v bloku nachádza, toľko násobne vyššiu šancu blok má na vybratie serverom. To znamená, že diamanty môžu zabezpečiť 1 až 64 násobne vyššiu šancu na vylosovanie. Diamanty reprezentujú výkonnostný index ťažiaceho počítača. Čím silnejší je počítač, tým väčšiu šancu na vyťaženie má.

Diagram spôsobu vyberania Bitcoin adresy, ktorá dostane odmenu v podobe Coinbase transakcie je možné vidieť na obrázku 6.

#### D. Blockchain Explorer Block

Blockchain Explorer poskytuje jednoduché listové zobrazenie všetkých transakcií, ktoré boli na serveri vytvorené. O jednotlivých transakciách sa zobrazujú nasledujúce informácie:

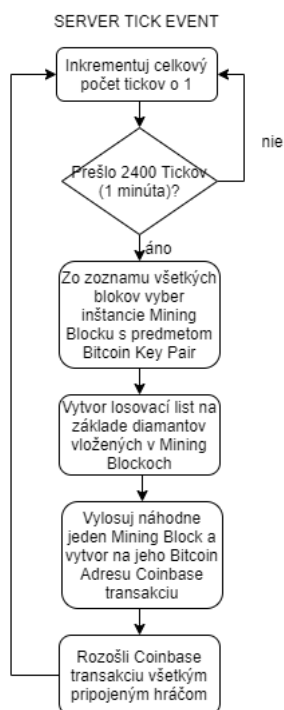
- Informácia či je transakcia používateľská alebo výsledok ťaženia (Coinbase)
- Adresa prijímateľa
- Množstvo posielaných Bitcoinov
- Adresa prijímateľa zvyšku (nepovinné)
- Množstvo zvyšných Bitcoinov (nepovinné)
- Informácia či sú výstupy už minuté alebo ešte nie

#### E. Ako modifikácia vyučuje Bitcoin

Pri skutočnom Bitcoine platí to, že aby užívateľ mohol sieť Bitcoin plnohodnotne používať, tak si potrebuje vygenerovať minimálne jeden privátny kľúč, z ktorého sa následne vygeneruje Bitcoin Adresa. Tieto kľúče si väčšinou užívateľ generuje pomocou Bitcoin peňaženiek. Keď má



Obr. 5 GUI Mining Blocku



Obr. 6 Diagram náhodného výberu Bitcoin Adresy ktorý dostane odmenu z ťažby

užívateľ vygenerovaný aspoň jeden pár kľúčov, tak môže prijímať platby, posilať platby alebo ťažiť Bitcoin. V rámci modifikácie platí to isté. Modifikácia hráčovi neumožní robiť nič do tej doby, než si nevygeneruje aspoň jeden predmet **Bitcoin Key Pair** v **Bitcoin Wallet Blocku**, ktorý v sebe nesie privátny kľúč a Bitcoin Adresu. Potom môže na takéto Bitcoin Adresy prijímať platby, odosielať platby, či ťažiť Bitcoin. Porovnanie kľúčových párov medzi skutočným Bitcoinom (rozhranie bitaddress.org) a modifikáciou je možné vidieť na obrázku 8.

Vytváranie platieb sa takisto rieši pomocou Bitcoin peňaženiek. Jediné, čo k tomu užívateľ potrebuje je mať privátny kľúč, ktorý má prístup k Bitcoinom na výstupoch transakcií a Bitcoin Adresu príjemcu. Potom môže užívateľ vytvoriť novú transakciu, pomocou ktorej minie Bitcoin, ku ktorým má prístup. Rovnako to platí aj v rámci tejto modifikácie. Hráč v rozhraní **Bitcoin Wallet Blocku** vloží do políčka predmet **Bitcoin Key Pair**, ktorý obsahuje Bitcoin Adresu, na ktorú sú vytvorené výstupy transakcií. Potom do prvého textového poľa zadá množstvo Bitcoinov, ktoré chce poslať a do druhého textového poľa zadá Bitcoin Adresu, ktorú si vypýta od hráča, ktorému chce Bitcoin poslať (prípadne použije Bitcoin Adresu z niektorého zo svojich ostatných predmetov Bitcoin Key Pair čím si hráč môže posilať Bitcoin aj medzi sebou). Následne



```
-----  
Transaction ID: 1  
Sender: 19a29756e  
Recipient: 11e60993d  
Amount: 1.0 Available  
  
Change: 19a29756e  
Amount: 49.0 Available  
  
-----  
Transaction ID: 2  
Coinbase Transaction  
Recipient: 1b1540bd1  
Amount: 50.0 Spent  
  
-----  
Transaction ID: 3  
Coinbase Transaction
```

Obr. 7 GUI Blockchain Explorer Bloku

tlačítkom Send transakciu vytvorí. Na obrázku 9 je možné vidieť porovnanie rozhraní pre tvorbu transakcií medzi Bitcoin peňaženkou Exodus a modifikáciou.

Blockchain je verejne dostupný a je možné ho prezerať pomocou lokálnych nástrojov v prípade, že má užívateľ stiahnutú kópiu Blockchainu u seba v počítači, alebo je možné k tomu využiť rôzne internetové prehliadače Blockchainu. V rámci Blockchainu môže užívateľ prezerať všetky transakcie, ktoré boli od počiatku siete Bitcoin až po súčasnosť vytvorené. Rovnakú možnosť ponúka aj modifikácia pomocou **Blockchain Explorer Bloku**. Umožňuje hráčom zobrazíť všetky transakcie vytvorené na serveri a k nim podstatné detaily. Jediný rozdiel je v tom, že modifikácia neorganizuje transakcie do blokov, ale iba do listu. Na obrázku 10 je možné vidieť porovnanie detailu transakcie skutočného Bitcoinu v rozhraní Blockchain.com a porovnanie detailu transakcie, ktoré ponúka modifikácia.

Ťažba sa v rámci Bitcoinu vykonáva ťažobnými softwami ako napríklad cgminer. Väčšinou sa jedná o jednoduché konzolové programy, ktoré na pozadí vykonávajú ťažbu a v rámci konzoly vypisujú základné informácie o ťažbe. Užívateľ si najskôr nastaví Bitcoin Adresu, na ktorú sa majú pripísať Bitcoin v prípade, že užívateľov počítač vyťaží blok a následne už iba zapne samotnú ťažbu. V rámci modifikácie **Mining Blocky** síce reprezentujú pomyselné počítače, ktoré "ťažia", no v skutočnosti tieto bloky na pozadí nerobia nič, iba server oboznamujú s Bitcoin Adresami, z ktorých vždy náhodne vyberie jednu, ktorej pripíše Bitcoin. Hráč do **Mining Bloku** vloží predmet **Bitcoin Key Pair** a o nič viac sa nestará. Keď chce hráč zvýšiť šancu Mining Bloku na vyťaženie Bitcoinov, tak doňho môže vložiť diamanty. Tieto diamanty symbolizujú výkon počítača. Čím výkonnejší počítač je, tým viac hashov dokáže za sekundu vyrábať a tým vyššiu šancu na vyťaženie bloku má. Z funkčného hľadiska sú ťažobné softwary a modifikácia úplne odlišné, avšak z užívateľského hľadiska sú takmer totožné.

#### F. Použitie modifikácie hráčmi

Keď hráči začnú hrať hru s nahranou modifikáciou, tak si v prvom rade musia byť schopní vytvoriť trojicu blokov, ktoré modifikácia do hry pridáva. To je možné vykonať dvomi spôsobmi. Výrobou, alebo administrátorskými príkazmi podľa toho, aké právomoci hráč na serveri má a akým spôsobom chce hrať. V prípade výroby sa jedná o nasledujúce tri výrobné recepty:

- **Bitcoin Wallet Block** - 2x papier
- **Bitcoin Mining Block** - 2x drevený krompáč
- **Blockchain Explorer Block** - 2x kniha

V prípade, že hráč má administrátorské práva, tak si bloky môže vyvolať nasledujúcimi príkazmi:

- /give @p bitcoinmod:wallet\_block
- /give @p bitcoinmod:mining\_block
- /give @p bitcoinmod:blockchain\_block

Každý hráč by si mal ako prvé vytvoriť **Bitcoin Wallet Block**. V rámci tohoto bloku si hráč môže vygenerovať predmet (alebo viacero predmetov) typu **Bitcoin Key Pair**, ktorý v sebe nesie pár privátneho kľúča a Bitcoin Adresy. Ukázaním myšou na tento predmet v inventárovom zobrazení sa ako tooltip ukáže informácia o privátnom kľúči a Bitcoin adrese.

Keď chce jeden hráč druhému poslať Bitcoin (prípadne keď si chce jeden hráč poslať Bitcoin v rámci svojich dvoch Bitcoin Adries), tak do rozhrania bloku **Bitcoin Wallet Block** vloží predmet **Bitcoin Key Pair**, ktorý obsahuje kľúčový pár, z ktorého chce Bitcoin poslať. Potom do prvého textového poľa v grafickom rozhraní napíše množstvo Bitcoinov, ktoré chce poslať a do druhého textového poľa napíše Bitcoin Adresu na ktorú chce Bitcoin poslať. Po stlačení tlačidla **Send** sa odošle požiadavka o novej transakcií na server, ktorý ju skontroluje a v prípade, že takúto transakciu je možné vytvoriť, tak ju server vytvorí a informuje o tom klientské aplikácie všetkých hráčov.

Keď sa hráč chce zaradiť do procesu ťaženia (miningu) a mať šancu na vyťaženie Bitcoinov, tak si potrebuje vytvoriť **Mining Block** (alebo viacero blokov). Do tohoto bloku hráč vloží predmet **Bitcoin Key Pair**, ktorý obsahuje Bitcoin Adresu, na ktorú chce hráč ťažiť Bitcoin. Mining Block, ktorý má v sebe vložený Bitcoin Key Pair má šancu, že ho server náhodne vyberie a na Bitcoin Adresu z vloženého predmetu Bitcoin Key Pair vygeneruje Coinbase transakciu. Čím viac Mining Blockov vo svete existuje, tým majú menšiu šancu, že ich server vyberie. Napríklad ak vo svete existuje 100 Mining Blockov, tak šanca, že nejaký Mining Block bude vybraný je 1:100. Každý hráč si môže vytvoriť ľubovoľný počet takýchto Mining Blockov.

K tomu, aby bola modifikácia a proces ťažby trošku interaktívnejší, tak hráči môžu vo svete ťažiť diamanty, ktoré sú pomerne vzácne. Tieto diamanty potom môžu vložiť do svojho Mining Blocku, čím zvýšia šancu, že server takýto blok vyberie.

Hráči majú prístup ku celej kópii herného Blockchainu pomocou **Blockchain Explorer Blocku**. Tento blok slúži iba pre informatívne účely, kde si hráči môžu prezrieť všetky vytvorené transakcie, výšky jednotlivých výstupov, príjemcov jednotlivých výstupov a stav (minutý/neminutý) jednotlivých výstupov.

Modifikácia do hry nepridáva žiadny spôsob, ktorým by sa dal získaný Bitcoin minúť (napríklad v nejakom obchode). Avšak táto modifikácia pridáva do hry ekosystém, ktorý hráči môžu využiť na obchodovanie medzi sebou.

Modifikácia zachováva pôvodný ekosystém Bitcoinu. To znamená, že maximálny počet Bitcoinov, ktoré modifikácia dokáže vyprodukovať je 21 miliónov. Počet Bitcoinov vypúšťaných pomocou Coinbase transakcií sa vždy po dosiahnutí určitej prahovej hodnoty zníži o polovicu. Počiatočná hodnota Coinbase transakcií je 50 Bitcoinov. Jediný rozdiel je ten, že modifikácia produkuje nové Bitcoinov každú minútu miesto každých 10 minút, aby to bolo zaujímavejšie pre hráčov.

#### *G. Porovnanie architektúr Bitcoinu a modifikácie do Minecraftu*

Vzhľadom k tomu, že Bitcoin funguje na architektúre P2P a Minecraft na architektúre klient-server, tak bolo potrebné niektoré koncepty pôvodnej siete Bitcoin pozmeniť tak, aby ich bolo možné simulovať na architektúre Minecraftu. V prvom rade zjednodušený Blockchain (TransactionsCapability) je spravovaný iba serverom, pretože server jediný si ukladá dáta do dlhodobej pamäti. Klientské aplikácie hráčov si taktiež vedú kópiu Blockchainu (v čase keď sú na server pripojené), avšak nijakým spôsobom tento Blockchain nespravujú, ale iba kopírujú zmeny, ktoré na ňom vykoná server. Blockchain na serveri slúži pre kontrolu a tvorbu nových transakcií a pre dlhodobé uloženie. Blockchain na klientských aplikáciách slúži pre zobrazenie zostatkov a pre prehliadanie transakcií v rozhraní Blockchain Explorer Blocku.

Taktiež sa musela pozmeniť aj ťažba. V skutočnej sieti Bitcoin sa do procesu ťažby zapájajú všetky Full-blockchain a Mining uzly a celý proces sa spolieha na peer-to-peer komunikáciu. Proces ťažby sa v sieti Bitcoin skladá z dvoch častí: kontrola transakcií a vyťaženie bloku, za čo uzol získa odmenu. V prípade tejto modifikácie kontrolu všetkých transakcií vykonáva iba server. Preto sa musela vymyslieť alternatíva, ktorá by hráčom umožnila získať Bitcoin. Preto sa do hry implementoval Mining Block, ktorým hráči na server zaregistrujú svoje Bitcoin Adresy, ktoré chcú zapojiť do ťažby a server z týchto adries náhodne každú minútu vyberie jednu adresu, na ktorú vygeneruje Coinbase transakciu. Tieto Mining Blocky aktívne nevykonávajú nič okrem toho, že server oboznamujú s Bitcoin Adresami hráčov a umožňujú im tak ťažiť Bitcoin. Ťažba síce na pozadí funguje inak, ale z používateľského hľadiska je rovnaká, pretože ani v skutočnej sieti Bitcoin užívateľ veľmi nevie, čo sa vykonáva na pozadí ťažiacieho softwaru.

Diagramy architektúr Bitcoinu a modifikácie je možné vidieť na obrázkoch 11 a 12 na konci tohoto článku. Tieto diagramy zobrazujú dôležité entity v rámci oboch technológií, údaje ktoré si uchovávajú, funkcie ktoré plnia a spôsob interakcie medzi sebou.

#### *H. Zhodnotenie výsledku*

Vytvorená modifikácia zjednodušene simuluje niektoré základné koncepty kryptomeny Bitcoin. Medzi tieto koncepty patrí generovanie párov privátneho a verejného kľúča, ktoré sa používajú

k dokázaniu vlastníctva k Bitcoinom a zároveň pre identifikáciu príjemcu transakcií. Následne modifikácia do hry prináša systém transakcií, ktoré fungujú podobným spôsobom ako v pôvodnej sieti Bitcoin, len trochu zjednodušene. Ďalej modifikácia do hry pridáva ťažbu, ktorá je ale prispôbená klient-server architektúre vzhľadom k tomu, že Bitcoin využíva peer-to-peer architektúru. Nakoniec modifikácia pridáva prehliadač zjednodušeného Blockchainu. Blockchain v prípade tejto modifikácie predstavuje iba list transakcií, neobsahuje organizáciu do blokov.

Do modifikácie sa na druhú stranu nedostali niektoré komplexnejšie koncepty Bitcoinu. V prvom rade to sú transakčné skripty, ktorých implementácia by bola zložitá a mohla by byť náplňou samostatnej práce z dôvodu rozsiahlosti problematiky. Ďalej to je organizácia transakcií do blokov, čo by bolo problematické z dôvodu obmedzenej serializácie a deserializácie dát v Minecrafte a z dôvodu obmedzených grafických rozhraní. Koncepty Bitcoinu bolo potrebné zjednodušiť aj z toho dôvodu, že modifikácia beží na serveri a nemôže tento server zbytočne spomaľovať, aby hráčom nekazila zážitok z hry.

Spojením Bitcoinu a Minecraftu sa vytvoril spôsob, ktorým sa dá Bitcoin vyskúšať a dajú sa o ňom získať základné znalosti jednoduchou a zábavnou formou bez potreby si študovať hlboké technické pozadie. Výhodou tohoto riešenia je, že užívatelia nemusia investovať skutočné peniaze k tomu, aby si Bitcoin odskúšali. Pretože aby sa človek dostal k skutočným Bitcoinom, tak si ich potrebuje od niekoho kúpiť osobne, kúpiť cez burzové stránky, alebo vyťažiť. Ťažba síce je zadarmo (ak sa neberie v úvahu cena elektriny), no kvôli veľkému počtu minerov je šanca na ich vyťaženie takmer nulová.

## VI. ZÁVER

Výsledkom tohto článku je modifikácia do hry Minecraft, ktorá do hry implementuje zjednodušený ekosystém Bitcoinu. Prináša do hry zjednodušenú funkcionálnosť Bitcoin peňaženiek, pseudo-mining a jednoduchý prehliadač "Blockchainu" (zoznamu transakcií). Toto všetko funguje ako v režime jedného hráča, tak aj na dedikovanom serveri v režime pre viacerých hráčov. V prípade záujmu o vyskúšanie modifikácie je možné si ju stiahnuť z [GitHub<sup>1</sup>](#).

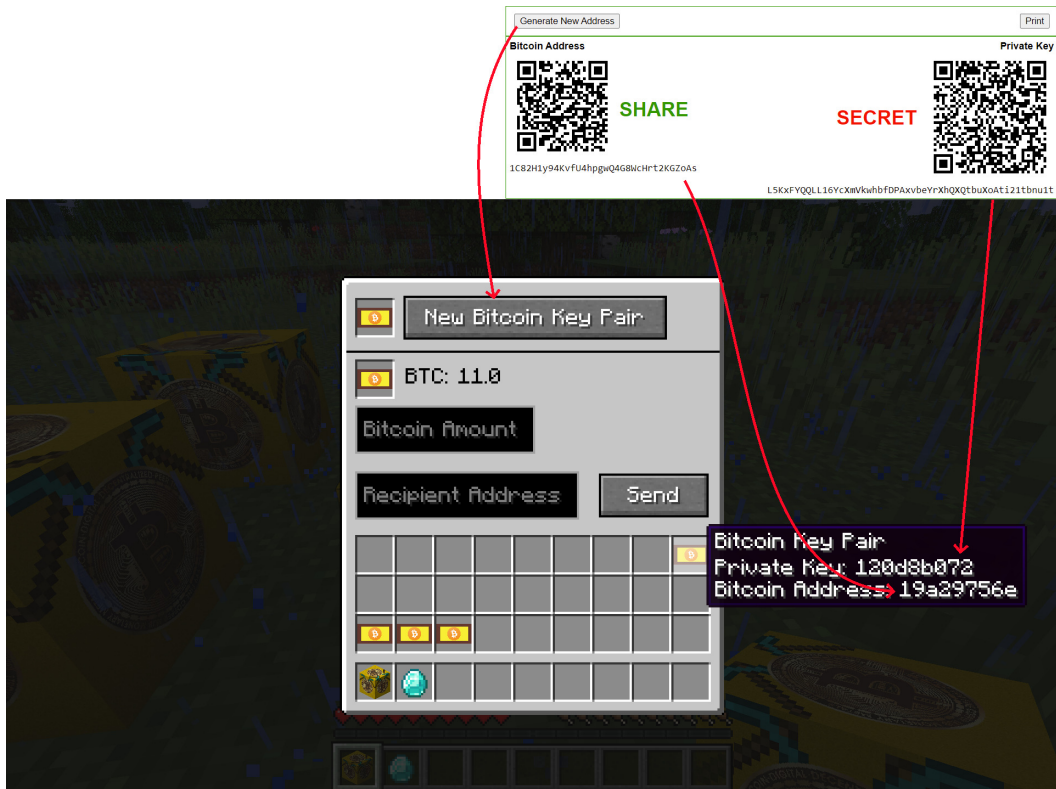
## VII. PODĀKOVANIE

Tento výskum bol podporený projektom 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market.

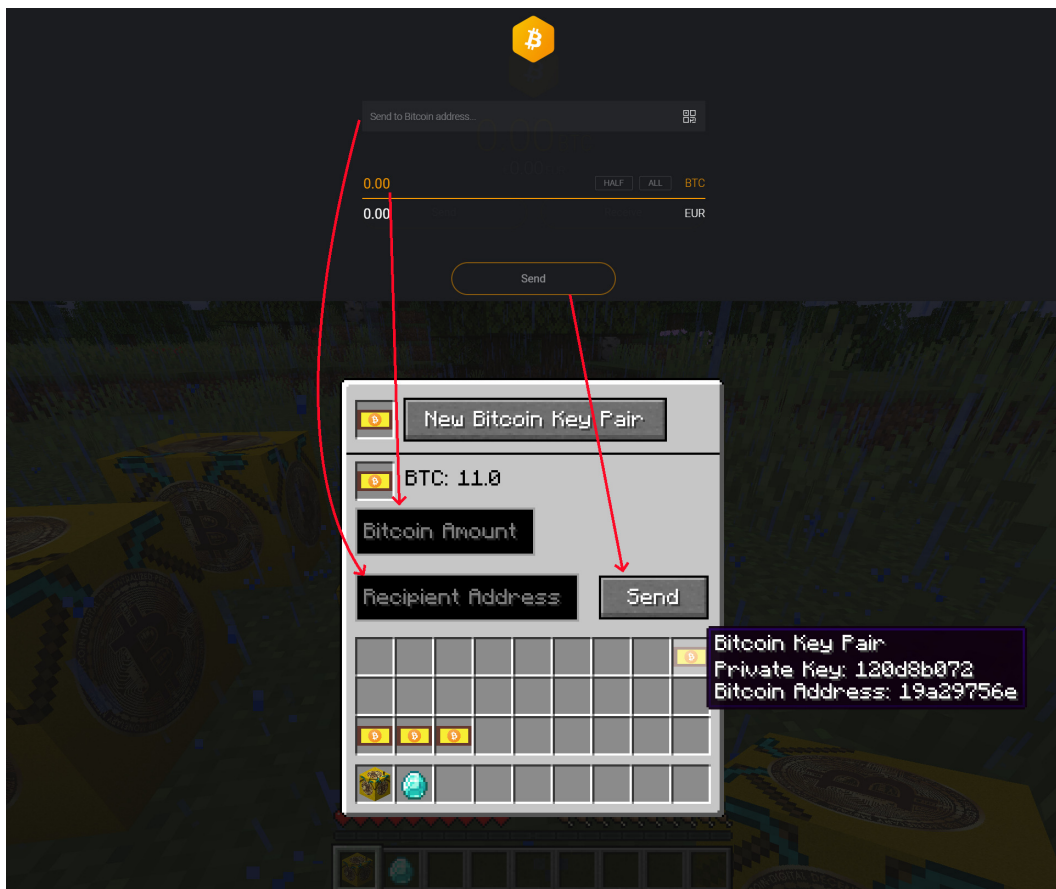
## LITERATÚRA

- [1] J. F. Buford, H. Yu, and E. K. Lua, *P2P Networking and Applications*. Morgan Kaufmann Publishers, 2009.
- [2] Andreas M. Antonopoulos, *Mastering Bitcoin*, 2nd ed. O'Reilly, 7 2017.
- [3] Wiki contributors, "Bitcoin Wiki," <https://en.bitcoin.it/>, 2021, [Online; navštívené 10-Apríl-2021].
- [4] Bitcoin.org, *Bitcoin Documentation*, <https://developer.bitcoin.org/>, 2021, [Online; navštívené 10-Apríl-2021].
- [5] MCForge, *Forge documentation*, <https://mcforge.readthedocs.io/en/1.15.x/>, 2021, [Online; navštívené 10-Apríl-2021].

<sup>1</sup><https://github.com/DanVUT/BitcoinModification>



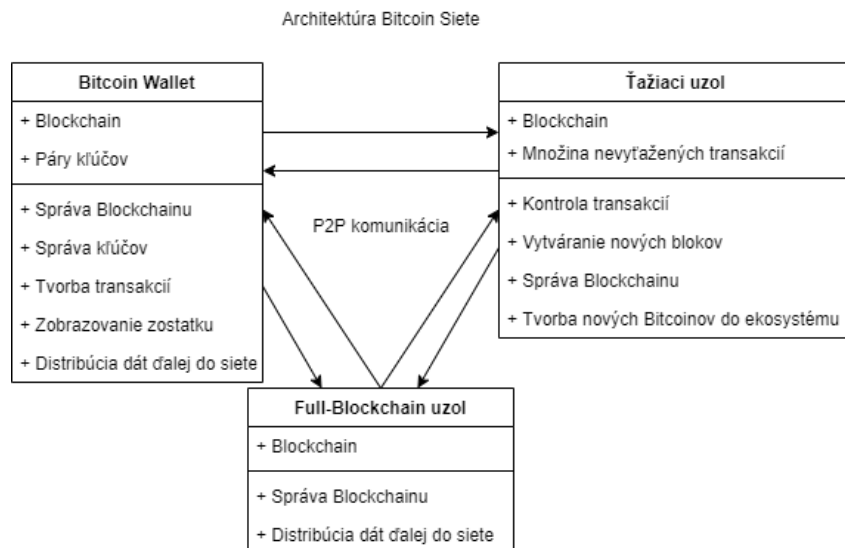
Obr. 8 Porovnanie kľúčových párov medzi skutočným Bitcoinom a modifikáciou



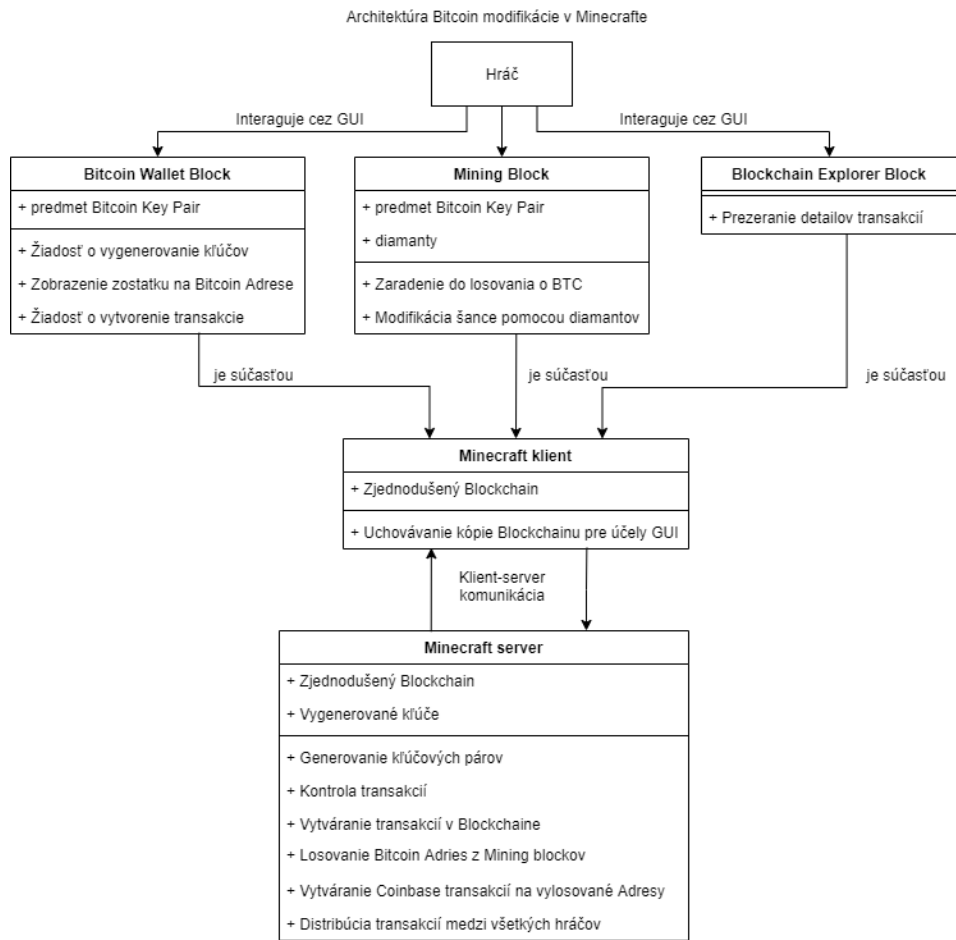
Obr. 9 Porovnanie vytvárania transakcií medzi peňaženkou Exodus a modifikáciou



Obr. 10 Porovnanie zobrazenia detailu transakcie medzi rozhraním Blockchain.com a modifikáciou



Obr. 11 Znážornenie architektúry Bitcoinu



Obr. 12 Znáznornenie architektúry modifikácie v Minecrafte

# The therapy of upper limbs with virtual reality technology

<sup>1</sup>Branislav SOBOTA, <sup>2</sup>Sára JAVORKOVÁ, <sup>3</sup>Marián HUDÁK, <sup>4</sup>Štefan KOREČKO

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

<sup>1</sup>branislav.sobota@tuke.sk, <sup>4</sup>stefan.korecko@tuke.sk

**Abstract** — This paper describes the therapy of upper limbs with virtual reality technology. Main focus of the paper is to the description of calculating bone rotation system. The user can add more upper limb virtual objects to the virtual environment and he can also add objects used in the system. The user can move these objects. If they are within range, the system calculates the animation method. The functions used in these calculations are useful for creating additional animations with various objects. With this system, the patient can be stimulated under the supervision of a therapist to practice some rehabilitation procedures. Due to the use of shared web reality, this process can be also applied in a distance form.

**Keywords** — upper limbs rehabilitation, inverse kinematics, A-frame, virtual reality

## I. INTRODUCTION

Virtual reality (VR) affects a person's perception and thus his nervous system. Development of VR and related technologies, in recent years, allows its application in areas where it was previously difficult or not possible at all, especially in human-machine interaction [1][2]. One such area is the area of neurorehabilitation. Many people have come to impaired motor pathways in the brain due to head injuries or diseases, leading to physical disabilities. One of the most common causes of physical disability for adults is stroke. A stroke involves the reduction of blood flow and insufficient oxygenation of the brain tissue, which can cause failure or loss of function of the affected brain area. This results often in an inability to move one or more limbs on one side of the body (hemiparesis). Also, a person may not be able to understand speech or speak, may have memory impairments, or loss of vision to one side. Several studies have shown that neurorehabilitation helps to achieve beneficial results in the treatment of patients with paresis. These studies also include the research of the Slovak Academy of Sciences (cooperating organization within the project solved in the authors' home laboratory LIRKIS). Technologies such as Mirror Boxing or robotic arm system were used in the SAS research together with EEG [1]. Until now, however, virtual reality technologies have not been used in this research. Thus, the question is whether it would be possible to positively influence the neurorehabilitation/therapy process by providing a suitable virtual environment.

## II. NEUROREHABILITATION AND VIRTUAL REALITY

In the field of medical applications, the VR and related technologies offers a variety of interaction techniques rendering effective for treatment and rehabilitation of human movement disorders [4] [5]. Upper limb rehabilitation process [6], practiced using head mounted displays with other equipment based on the rehabilitation needs [7] can serve as a commendable example. In a virtual environment, it is possible to simulate different conditions of a virtual hand, visualize different situations for the user and assign them tasks to perform utilizing virtual hands [8], [9]. Due to the pandemic situation, it is difficult to perform personal rehabilitation and training of patients [10]. Novel science-based treatments are considering remote therapy options, where a therapist can communicate with a patient and oversee the training process [11]. In order to the standard VR applications, web-based VR can provide sharing of virtual content and user collaboration in the real-time [12]. A widely used rehabilitation application is a MS Kinect based rehabilitation called *Removiem* [13]. It is created for people with multiple sclerosis who need to practice the movement of the whole body. The application offers several exercises that the patient

can practice on their own. The patient needs a Kinect device and a computer with a monitor, which shows a virtual environment for exercise. In addition to the fact that the patient can exercise without the therapist's supervision, the therapist is able to control his or her results and adapt the program plan to the patient. Another example is the *IREX* (Interactive Rehabilitation and Exercise Systems) project from GestureTek [14]. Unlike the previous one, this project seeks to disconnect people from the reality of lengthy and boring rehabilitation. Each part of rehabilitation is given in the form of simple games (gamification), in which both children and adults have fun. The application contains several games that deal with balancing, rotating, bending or lifting things.

Health problems caused by stroke include inability to move one or more limbs on one side of the body (hemiparesis) [15]. Neurorehabilitation is one of the main therapies of stroke treatment. The basic pillar used in neurorehabilitation is neuroplasticity, which is brain property. Due to this property, it is possible that after damage to the brain area, another area that was not originally used for the given function will partially take over the function of this area.

Slovak Academy of Sciences - SAS implemented a BCI system with a robotic arm and used them in its research of the neurorehabilitation for people affected by upper limb paresis [15]. The goal is to aid with patient movement through the robotic arm. Neurorehabilitation consists of the functional replacement of a damaged brain area with another area of the brain that was not originally intended for it. This goal cannot be achieved with classical rehabilitation, where the movement of the affected limb is performed by a therapist. In robot-assisted rehabilitation, the movement of the limb is performed by a robotic arm. The robotic arm individually and optimally adjusts the training and evaluates the progress of the patients based on results from EEG.

Another technique SAS used in the research is rehabilitation using motoric projection (MP). The patient has the idea of the movement of his limb, but the movement itself is not performed. Although the combination of MP and classical rehabilitation is more effective, it also has its disadvantages. The movement of a given limb performed by a therapist does not have to correspond with the patient's intention.

The project used a developed BCI system based on motor resonance scanned by EEG. The patient tried to move using an MP and a robotic arm. The goal of the MP-BCI system is to use patient mental activity associated with the movement of the robotic arm and thus return him the lost ability to move the limb. However, it requires a great concentration on the activity.

For this research, a virtual reality system was created to help with neurorehabilitation of a patient with upper limb paresis. It is based on the LIRKIS G-CVE (Global Collaborative Virtual Environments, author's home laboratory development) [16]. The system was divided into a subsystem for the patient and a subsystem for the therapist. An important part of the application is the simulation and visualization of the movement of the patient's upper limb (Fig. 1), including the grips of various objects. In this case, we could work with animations in two ways. The first is most often used in cases, when the animation does not need to change, for example simple walk animation, which is still the same and does not need to adapt to environment. Second way is to import object with already created bones and it is easily manageable. It is needed to create animations, which can adapt and catch target object, even when it is moved in different place. The first way is insufficient for this type of implementation. A short description of second way implementation using A-frame and/or Three.js frameworks is in follows.

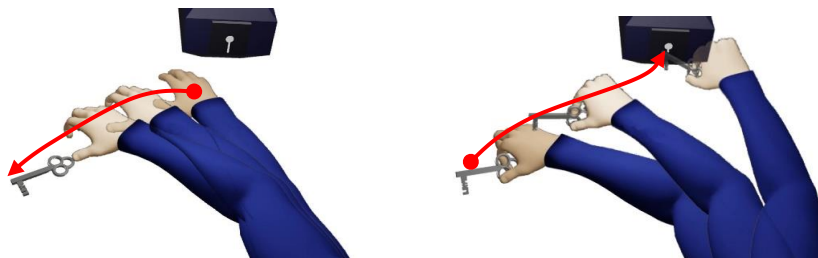


Fig. 1 An examples of upper limb movement animation in virtual reality (patient's view)

### III. CALCULATION BONE ANGLES

A model of an arm with bones exported from Blender application as *.gltf* file was used. A-frame can recognize bones and manage them. Every bone has axis on its own according to the previous one. The position of bones is not recommended to change, because previous bone does not change with it. As in the real world, human needs to change rotation of previous bones to get desirable position of target bone. The same principle is used in inverse kinematics. As an example for this article, was chosen animation with keys and locks. The goal of animation is hand to catch the key, move it into the lock, rotate it in and return to the starting position.



### A. Inverse kinematics

The problem, that inverse kinematics is solving, is to find out rotations of every part of object to reach some point in environment. Inverse kinematics is mostly used with robots with more joints to reach some point [17]. It is acceptable for robots to move in each direction, but not with human's arm, which has limits. The biggest issue with normal calculation of rotations is that there is a lot of possible solutions for each bone. At the end of the calculations, rotations can be unnatural for human body.

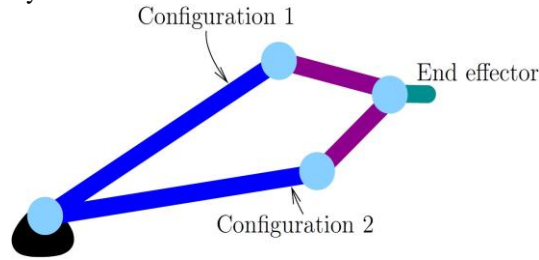


Fig. 2 Inverse kinematics in 2D with 2 joints [18]

### B. Angle calculation

A point of environment to reach (key) and current rotations of every bone in arm are defined. It is sufficient to make calculation with only two joints of arm. Basic joint represents shoulder, which in these animations does not move, only rotate. Next joint represents elbow, which moves because rotation of previous bone changes and rotates itself to change position of a wrist, last joint in calculation. In this case, all we must do is to find out the location of wrist in environment. Unfortunately, there are infinite possibilities, where it can be. To be specific, there are whole circle of potential positions. Therefore, a little different solution was used.

The principle of this solution lies in the division of the axes into separate ones. Then the axis to zero value is reset. Simply, we put all the points into 2D space for each axis, except one. It is axis represents itself bone's rotation (in appropriate node) and it has no effect to position of next bone.

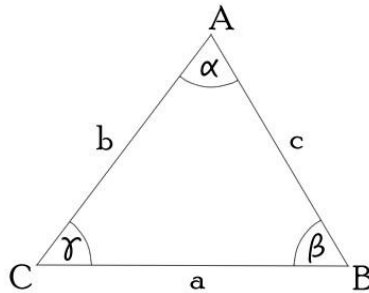


Fig. 3 Calculation of angles in 2D triangle

An ordinary 2D triangle is thus defined (Fig. 3). It is easy to calculate each angle in this triangle using the following formula:

$$\beta' = \arccos\left(\frac{a^2 + c^2 - b^2}{2ac}\right) \quad (1)$$

where:

- a - distance between shoulder's position (B) and position of key (C);
- b - distance between elbow's position (A) and position of wrist (C);
- c - distance between shoulder's position (B) and position of elbow (A).

The distance can be easily obtained using the *THREE.js* function *Vector3.distanceTo(Vector3)*, where the vectors are the positions of the selected bones.

When the calculation is done in all three axes, the rotations of the appropriate triangles (meaning: the hand holds the key) are obtained, but the results need to be adapted to the actual rotations of the arm. Therefore, the same calculations are needed, but in a slightly different triangle, to find the differences between the actual and desired rotations. It was only necessary to change one variable: b - distance between elbow's position (A) and position of key (C);

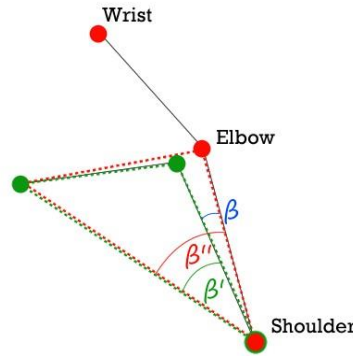


Fig. 4 Calculation of desired rotations for shoulder

As is shown in Fig. 3, green triangle is desired one, which is already defined from previous calculation. The red one is representing actual positions of bones. These two triangles were used to find the angle at which the bone rotates. After inserting the results into the following formula (2), the rotations of the arm were obtained in the actual state:

$$\beta = \beta'' - \beta' \quad (2)$$

For calculations of elbow's rotation similar principle was used. The same desired triangle was used. Another angle in it was calculated only. As is shown in the Fig. 5, it is needed to reduce the angle from 180° to get the desired angle.

$$\alpha = 180 - \arccos\left(\frac{a^2 + b^2 - c^2}{2ab}\right) \quad (3)$$

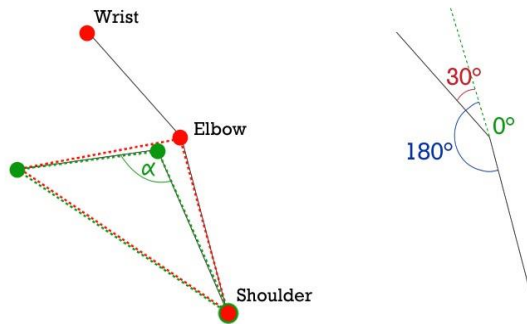


Fig. 5 Calculation of desired angle for elbow

Finally, it should be noted that this calculation calculates the wrist position equal to the key. But that was not our intention. In the real world, when a hand has to pick up a key, the target position of the hand is slightly above the key. All that needed to be done was to change the target position slightly higher in the Y-axis.

### C. Pick up the target

The rotation of the fingers is set statically to the key lifting position. In this part, we focused mainly on the rotations and positions of the key. The key was added as a child to the thumb. This means that the world given to the descendant depends on the parent. If a parent is moved then descendant would move with it at the same time. It is not necessary to constantly change position of all the descendants, their position and rotation is calculated from the position of the parent as the base of the object world. There are two ways to assign a descendant to a parent, as is shown in the Fig. 6:

- *Object3D.add(Object3D)* – its local position stays the same, so it moves to the position based on new parent
- *Object3D.attach(Object3D)* – its global position does not change and local position is calculated based on new parent, so it looks like nothing changes on the outside

The first method was implemented, because it was needed to calculate position of the key one way or another.

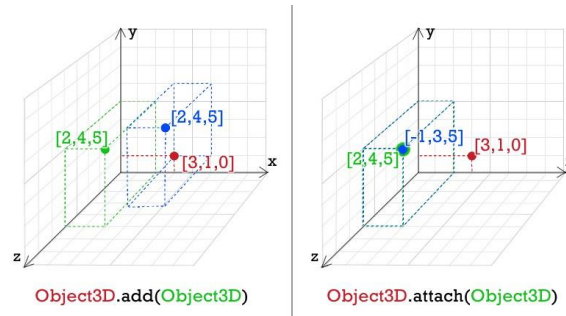


Fig. 6 Difference between add() and attach() methods

#### D. Movement to the lock

The same principle as in the next section was implemented for the movement to the lock and into the lock, just with different target. A rotation of the key aimed at the center of the lock for more realistic impression of animation was added.

#### E. Rotation around pivot

In the Fig. 7 is shown the principle of rotation around a pivot other than the center of the object, using the following entities (implementation as a cubes):

1. lock center of gravity;
  2. the center of gravity of the wrist transferred to the local space of the lock with the zero X-axis;
  3. target position of the 2nd cube;
  4. position of the 2nd cube transferred to the local space of the wrist with zero X-axis;
  5. the center of gravity of the wrist before moving the wrist to the 4th cube position.
- wherein the 2nd and 3rd cubes are descendants of the 1st cube and the 4th cube is a descendant of the 5th cube.

For this calculation, it is important that the cube whose descendants. From the first three cubes, it is possible to calculate the angle by which the hand should turn around the lock. The cubes 1 and 5 will rotate at the same angle simultaneously with the rotation of the wrist on the Y-axis. Their descendants will also rotate with them. The position of the wrist is moved to the new position of the 4th cube with each change. Since the key is still a descendant of the wrist, it will rotate with it.

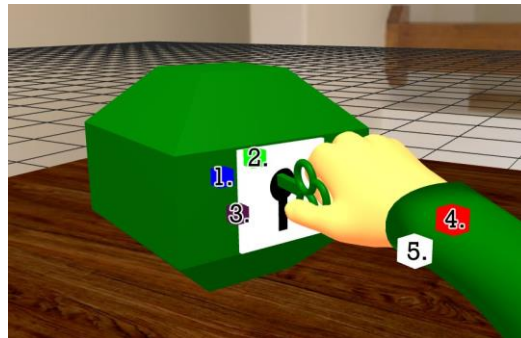


Fig. 7 Rotation according to the specified pivot

## IV. ANIMATION MANAGEMENT

A rotations at the end of each part of animation were calculated in previous section. It was needed to get the bones and the objects to the target coordinates. A system of components in *A-frame* was used, which contains the function called *tick()*. This function is called in every frame, so the rotation of the bones is continuous in this method (not in method using clock measurements).

For performance optimization, we decided to calculate how many frames each animation part takes. Thus, it was possible to calculate the rotation for one frame. This value is based on frames per second in the currently used device. To detect rotation in one frame, rotation at the beginning and end of the animation part was needed.

$$oneFrameAngle = \frac{startedAngle - endedAngle}{fps} \quad (4)$$

When one part of animation is done then next one follows till the end of whole animation. An enumeration type was used in management of the animation part. This contains all parts and one variable that stores the actual one. Almost all parts of animations are calculated on this base, but rotation in a lock is calculated in real time according to the methods in previous section, because just the position at the end of part is not enough.

## V. CONCLUSION

The described system of calculating bone rotation is a solution to several problems. The user can add more upper limb objects to the environment and can also add objects used in the system (for example, more keys and locks or dice or cups, etc.). The user can move objects. If they are within range, the system calculates the animation method. The functions used in these calculations are useful for creating additional animations with various objects.

The whole system allows you to work with either the right or left upper limb or both limbs at once in virtual environment. However, at this level of therapy, coordination of the two upper limbs would be difficult for patients and is only performed once the principal control of the upper limbs has been restored. Thus, at this level of therapy, the dominance of an upper limb is not significant (dominance is given by the patient's previous habit) and the main task is to monitor and perform therapy between the brain and the limb/s in the most basic activities. As was mentioned in introductory subchapters described bone calculation system was implemented in the application for the therapy of upper limbs with virtual reality technology. After completing the rehabilitation parameters, the patient is notified that the training will begin. For training purposes, it is essential that the patient sits at the table and perceives the 3D hand model (Fig. 8). After starting the training, several hand animations are performed. If the patient feels uncomfortable, the therapy process is stopped and then continues later. Given that each patient can perceive the 3D hand differently, we decided to extend the interface with user profiles. These profiles store data on training, hand configuration in space, and its color.

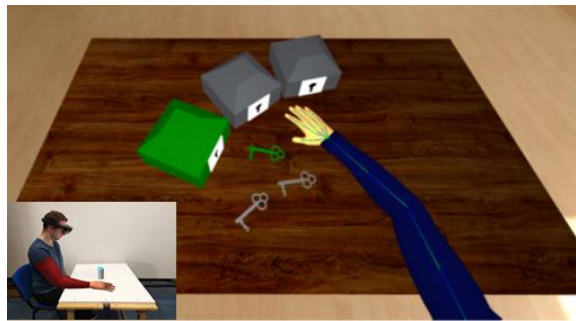


Fig. 8 Real therapy scene: with patient

## ACKNOWLEDGMENT

This work has been supported by the APVV grant no. APVV-16-0202 “Enhancing cognition and motor rehabilitation using mixed reality”.

## REFERENCES

- [1] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, and J. Gain, “A survey of augmented, virtual, and mixed reality for cultural heritage,” *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 11, no. 2, pp. 1–36, 2018.
- [2] A. P. R'ios, V. Callaghan, M. Gardner, and M. J. Alhaddad, “Using mixed-reality to develop smart environments,” pp. 182–189, 2014.
- [3] R. Rosipal. Brain-computer interface with robot-assisted training for rehabilitation (BCI-RAS). url: <http://aiolos.um.savba.sk/~roman/rrLab/projects.html>.
- [4] D. Sapkaroski, M. Baird, J. McInerney, and M. R. Dimmock, “The im-plementation of a haptic feedback virtual reality simulation clinic withdynamic patient interaction and communication for medical imaging stu-dents,” *Journal of medical radiation sciences*, vol. 65, no. 3, pp. 218–225, 2018.
- [5] D. S. Lopes and J. A. Jorge, “Extending medical interfaces towards vir-tual reality and augmented reality,” *Annals of Medicine*, vol. 51, no. sup1, pp. 29–29, 2019.
- [6] C. Nissler, M. Nowak, M. Connan, S. B'uttner, J. Vogel, I. Kossyk, Z.-C.M'arton, and C. Castellini, “Vita—an everyday virtual reality setup forprosthetics and upper-limb rehabilitation,” *Journal of neural engineering*, vol. 16, no. 2, p. 026039, 2019.
- [7] D. Chen, H. Liu, and Z. Ren, “Application of wearable device htc vive inupper limb rehabilitation training,” pp. 1460–1464, 2018.
- [8] X. Wu, H. Liu, J. Zhang, and W. Chen, “Virtual reality training systemfor upper limb rehabilitation,” pp. 1969–1974, 2019.
- [9] A. Elor, M. Teodorescu, and S. Kurniawan, “Project star catcher: A novel immersive virtual reality experience for upper limb rehabilitation,” *ACMTransactions on Accessible Computing (TACCESS)*, vol. 11, no. 4, pp. 1–25, 2018.

- [10] S. Becker, M. Chaple, T. Freese, H. Hagle, M. Henry, I. Koutsenok, L. Krom, R. Martin, T. Molfenter, K. Powell, et al., "Virtual reality for behavioral health workforce development in the era of covid-19," *Journal of Substance Abuse Treatment*, p. 108157, 2020.
- [11] S. Pedram, S. Palmisano, P. Perez, R. Mursic, and M. Farrelly, "Examining the potential of virtual reality to deliver remote rehabilitation," *Computers in Human Behavior*, vol. 105, p. 106223, 2020.
- [12] D. Seo, B. Yoo, and H. Ko, "Webizing collaborative interaction space for cross reality with various human interface devices," pp. 1–8, 2018.
- [13] J A Gil-Gómez, J A Lozano-Quilis, H Gil-Gómez. Virtual Reality System for Multiple Sclerosis Rehabilitation using KINECT. cit.: 05/2021, url: <http://rehab-workshop.org/2013/papers/136696042743759.pdf>.
- [14] GestureTek. "Irex is gesturetek health's flagship product for the healthcare arena." Cit.: 06/2021, url: <http://www.gesturetekhealth.com/products/irex>.
- [15] B. Cimrová, I. Farkaš, and R. Rosipal. "Využitie rozhrania mozog-počítač pri neurorehabilitácii: prehľad aktuálneho výskumu." *Kognícia a umelý život* 14 (2014).
- [16] M., Hudák, M., Sivý. Web-based collaborative virtual environments to support cross-platform access, In: *Poster 2019 International student scientific conference*, Prague, 2019, pp.178-182.
- [17] DeMers D., Kreutz-Delgado K.: "4 - inverse kinematics of dextrous manipulators," in *Neural Systems for Robotics*, O. Omidvar and P. van der Smagt, Eds. Boston: Academic Press, 1997, pp. 75–116. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780080925097500087>
- [18] Bower T.: Inverse kinematics. [Online]. Available: [http://faculty.salina.k-state.edu/tim/robotics\\_sg/Arm\\_robots/inverseKin.html](http://faculty.salina.k-state.edu/tim/robotics_sg/Arm_robots/inverseKin.html)

# Návrh grafického súborového formátu pre ukladanie dvoj-dimenzionálnych dát s využitím HDS

<sup>1</sup>Andrea PETRÁŠOVÁ, <sup>2</sup>Branislav MADOŠ, <sup>3</sup>Anton BALÁŽ

<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>andrea.petrasova@student.tuke.sk, <sup>2</sup>branislav.mados@tuke.sk, <sup>3</sup>anton.balaz@tuke.sk

**Abstrakt** — Článok sa zaoberá návrhom grafického súborového formátu pre uloženie dvoj-dimenzionálnej rastrovej obrazovej informácie s využitím bezstratovej kompresie dát založenej na použití hierarchických dátových štruktúr reprezentujúcich informáciu pomocou dvoj-dimenzionálnych ekvivalentov oktantových stromov resp. z nich odvodených orientovaných acyklických grafov. Navrhnutý súborový formát je orientovaný na ukladanie obrázkov s obmedzenou paletou farieb, kde sa predpokladá existencia väčších oblastí vyplnených obrazovými elementmi so zhodnými atribútmi. V poslednej časti článku sú zhrnuté výsledky testovania použitého kompresného algoritmu pri ukladaní dvoj-dimenzionálnej obrazovej informácie pri použití rôzneho počtu odtieňov sivej.

**Keywords** – *Sparse Voxel Octrees, SVO, Pointerless Sparse Voxel Octrees, PSVO, hierarchické dátové štruktúry, grafický súborový formát, Quadtree Image File Format, QIFF*

## I. ÚVOD

Množstvo dát, ktoré je každoročne vyprodukované technológiami určenými pre spracovávanie informácií a pre komunikovanie – Informačnými a Komunikačnými Technológiami (IKT) – angl. Information and Communication Technology (ICT), je obrovské a neustále sa zväčšuje.

Dáta sú predstavované e-mailovou komunikáciou, komunikáciou cez sociálne siete, video-konferenčnými hovormi, ale aj sms správami apod., keď časť týchto dát sa vygeneruje, prenesie a zlikviduje, ale časť digitálnej komunikácie zostane archivovaná a prípadne sa stane súčasťou záloh dát jednotlivých informačných systémov. Množstvo, predovšetkým multimediálnych dát, existuje v obrovskom počte kópií, keď napríklad Blu-ray film, zaznamenaný vo forme niekoľkých GB až desiatok GB dát, môže byť vydaný v náklade miliónov kópií, ktoré potom tvoria súčasť multimediálnych archívov. Významná veľkosť dát je predstavovaná počítačovými hrami. Veľké objemy dát tvoria tiež webové stránky publikované na internete, články publikované spravodajskými portálmi, rôzne digitálne archívy fotografií a ďalších multimédií. Dáta vznikajú aj ako súčasť aktivity odvetví ako je priemysel, kde slúžia pre účely riadenia a kontroly výrobných procesov, alebo ako výsledok diagnostických procesov v medicíne, kde vznikajú digitalizované záznamy z elektrokardiografu (EKG), alebo výstupy počítačovej tomografie Computed Tomography (CT), alebo magnetickej rezonancie – Magnetic Resonance Imaging (MRI). Objem produkovaných, generovaných a archivovaných dát sa neustále zvyšuje, jednak tým, že stúpa počet používateľov služieb a rozsah rôznych služieb, ktoré využívajú a jednak vzostupom náročnosti týchto služieb na množstvo dát. Príkladom môže byť zvyšovanie kvality filmov, keď sa zväčšenie rozlíšenia filmov z FullHD kvality do kvality 4K prejavuje štvornásobne väčším počtom bodov obrazu, čo má vplyv na to, koľko zaberá príslušný film priestoru na sekundárnom úložnom zariadení počítača, prípadne výmennom médiu pre takúto zariadenie, ako je napríklad Blu-Ray disk.

Aj keď sa kapacity pre prenos a uchovávanie dát neustále zväčšujú, je dôležité venovať pozornosť aj optimalizácii týchto dát, teda zabezpečeniu, aby sa negenerovali, neprenášali a najmä neuchovávali nad rámec optimálneho množstva. To je jeden z dôvodov, ktorý predstavuje motiváciu pre preskúmanie možnosti vytvoriť grafický súborový formát umožňujúci ukladanie dvoj-dimenzionálnych dát s využitím doménovo-špecifického kompresného algoritmu umožňujúceho bezstratovú kompresiu dát s využitím hierarchických dátových štruktúr založených na oktantových stromoch, resp. orientovaných acyklických grafoch.

## II. SÚVISIACE PRÁCE

Hierarchické dátové štruktúry, angl. Hierarchical Data Structures (HDS) sú používané pre reprezentáciu multi-dimenzionálnych dát už desiatky rokov, a to v podobe kvadrantových stromov (quadrees) pre dvoj-dimenzionálne dáta a v podobe oktantových stromov (octrees) pre troj-dimenzionálne dáta. Tak z kvadrantových stromov, ako aj z oktantových stromov, ktoré neumožňujú spájanie spoločných podstromov - common subtree merge (CSM), sú odvodené orientované acyklické grafy, ktoré toto spájanie spoločných podstromov umožňujú.

V 70-tych rokoch minulého storočia vzniklo viacero prác, venovaných tejto problematike, pričom medzi ne patria napríklad práce Gargantini [1], Hunter a Steiglitz [2], Klinger and Dyer [3]. V 80-tych rokoch minulého storočia to boli práce Kawaguchi a Endo [4], Srihari [5], Rubin a Whitted [6], Jackins a Tanimoto [7], Webber a Dillencourt [8], Samet [9] a Meagher [10][11][12].

Efficient Sparse Voxel Octrees (ESVO) zavedené v [13], sú hierarchickou dátovou štruktúrou, ktorá umožňuje odstránenie prázdnych podstromov. ESVO je na rozdiel od iných hierarchických dátových štruktúr schopný nahradiť celé podstromy špeciálnou kontúrovou informáciou a tak zvýšiť geometrické rozlíšenie a prináša kompaktnejšie zobrazenie hladkých povrchov. Kontúrová informácia má 32 bitov, keď 8 bitov tvorí kontúrová maska a 24 bitov tvorí ukazovateľ.

High Resolution Sparse Voxel Directed Acyclic Graphs sú orientované acyklické grafy, ktoré umožňujú spájanie spoločných podstromov resp. podgrafov a boli predstavené v roku 2013 v [14]. Maska dcérskych uzlov CHNM je 8b a je doplnená 24b, ktoré sú rezervované. Každý ukazovateľ na dcérsky uzol je 32b, takže všetky časti uzla, ako aj celé uzly a teda aj celá HR SVDAG HDS je zarovnávaná na 32b.

Symmetry-aware Sparse Voxel Directed Acyclic Graphs (SSVDAG) je hierarchická dátová štruktúra, ktorá má jeden druh uzlov, avšak jej listovú vrstvu tvorí vrstva trojdimenzionálnych mriežok voxelov s rozmerom  $4^3$  voxelov. SSVDAG majú implementované ukazovatele na dcérske uzly, pričom tieto môžu mať dĺžku 16b, 32b alebo 33b a obsahujú aj informáciu o možnom zrkadlení uzlov pri využívaní spájania spoločných podstromov, čo zvyšuje úroveň kompresie dát. Predstavená bola v roku 2016 v [15].

Pointerless Sparse Voxel Directed Acyclic Graphs (PSVDAG) je hierarchická dátová štruktúra určená pre reprezentáciu geometrie trojdimenzionálnych scén, navrhnutá v roku 2020 v [16]. PSVDAG v sebe spája výhody tak PSVO ako aj SVDAG HDS. Výhodou PSVDAG je, že nepotrebuje reprezentovať ukazovatele na dcérske uzly, čo znamená značnú úsporu veľkosti binárnej reprezentácie interného uzla. Týmto sa približuje svojou konštrukciou k PSVO. Na druhej strane, vďaka aplikovaniu systému návěstí a ich volaní umožňuje aplikovanie spájania spoločných podstromov, ktoré je charakteristické pre SVDAG.

V roku 2021 bola v [17] predstavená hierarchická dátová štruktúra založená na princípe PSVO, ktorá umožňuje reprezentovať nielen dva symboly, pre aktívne a pasívne voxely v rámci reprezentácie geometrie scény, ale viacero rôznych symbolov, ktoré umožňujú reprezentovať viaceré hodnoty príslušného atribútu obrazových elementov, ako je napríklad farba voxelu.

## III. NÁVRH GRAFICKÉHO SÚBOROVÉHO FORMÁTU

Navrhnutý grafický súborový formát sa skladá z troch častí, pričom prvé dve predstavujú metadáta, posledná časť predstavuje samotné obrazové dáta.

Prvou je hlavička súboru HEAD, ktorá sa skladá konštantne z 21 bajtov. Nasleduje voliteľná paleta farieb ICP a napokon nasledujú samotné dáta obrázka v podobe hierarchickej dátovej štruktúry, označené ako IMG. Každá z týchto troch častí súboru je zarovnaná na celé bajty, hoci tak paleta farieb ako aj samotné dáta obrázku nemusia byť po zakódovaní prirodzene zarovnané na celé bajty. Backus-Naurova Forma (BNF) reprezentujúca štruktúru navrhnutého súboru je:

QUIFF::=< HEAD >< ICP >< IMG >|< HEAD >< IMG > (1)

Kde platí nasledujúce:

<SYM> - povinný neterminálny symbol SYM,  
"sym" - terminálny symbol sym,  
(n) <SYM> - symbol SYM, konkatenovaný n krát,  
| - alternatívna,  
Juxtapozícia - zret'azenie

### A. Hlavička súboru HEAD

Prvých 9 bajtov predstavuje identifikátor súboru a pre túto iniciálnu verziu súborového formátu má hodnotu „QIFF2021a“. Nasleduje informácia o veľkosti ukladaného obrázku v počte pixelov v ose x a v ose y. Pre os x je táto informácia označená ako SIZE<sub>x</sub> zapísaná na 4 bajtoch, podobne pre os y je táto informácia označená ako SIZE<sub>y</sub> a je zapísaná tiež na 4 bajtoch.

Ďalej je v 1 bajte zapísaná informácia označená ako COL, ktorá hovorí o tom, či bude použitá paleta farieb alebo nie a či bude pre reprezentáciu farieb použité kódovanie s fixným počtom bitov na farbu, alebo s variabilným počtom bitov – Huffmanovo kódovanie.

Bez ohľadu na to, či je použitá farebná paleta alebo nie, súčasťou hlavičky je informácia COLSIZE, ktorá na 8 bitoch zaznamenáva, na koľkých bitoch je zapísaná konkrétna farba. Prípustné hodnoty sú z rozmedzia  $< 1; 32 >$ .

Poslednou informáciou je počet farieb farebnej palety COLNUM, zapísanou na 16 bitoch. maximálny počet farieb je preto 65536 (hodnota je do súboru zapísaná ako dekrementovaná o hodnotu 1). BNF forma hlavičky je zapísaná ako:

```
HEAD ::= < ID > < IMGS > < COL > < ICPS >
ID ::= "QIFF2021a"
IMGS ::= < SIZEx > < SIZEy >
COL ::= "00000000"|"00000001"|"00000010"|"00000011"
ICPS ::= < COLSIZE > < COLNUM >
SIZEx ::= (32) < BIT >
SIZEy ::= (32) < BIT >
COLSIZE ::= (8) < BIT >
COLNUM ::= (16) < BIT >
BIT ::= "0"|"1"
```

### B. Paleta farieb ICP

Indexovaná paleta farieb ICP môže ale nemusí byť v súbore prítomná, čo je závislé od toho, ako je nastavený parameter COL v hlavičke súboru. Ak je indikované, že paleta farieb bude súčasťou súboru s fixnou dĺžkou kódov farieb, v tom prípade je pre ňu vyčlenené miesto s veľkosťou ICP<sub>size</sub>. Je zapísaná ako sekvencia kódov farieb z indexovanej farebnej palety, pričom každý kód COLCODE je reprezentovaný sekvenciou COL<sub>size</sub> bitov a ich počet je COL<sub>num</sub>. BNF formálna reprezentácia palety farieb je nasledovná:

```
ICP ::= < COLS > < RES >
COLS ::= (COLNUM) < COLCODE >
COLCODE ::= (COLSIZE) < BIT >
RES ::= (0)*(7) < BIT >
BIT ::= "0"|"1"
```

### C. Obrazové dáta IMG

Vlastné obrazové dáta sú uložené vo forme hierarchickej dátovej štruktúry založenej na Pointerless Sparse Voxel Octrees dátovej štruktúre, bližšie popísanej v [17]. Tá umožňuje zapisovať viacbitové atribúty obrazových elementov, ktoré môžu byť interpretované ako farba obrazového bodu. Formálna reprezentácia tejto dátovej štruktúry vo forme BNF je:

```
PSVO ::= < INODE > | < LNODE >
INODE ::= < CHNM >
CHNM ::= (n) < HT >
HT ::= "0" < INODE > | "0" < LNODE > | "1" < SYMB >
LNODE ::= (n) < SYMB >
SYMB ::= (x) < BIT >
< BIT > ::= "0"|"1"
```

Kódovanie PSVO použitého v tomto grafickom formáte je také, že bit 0 je použitý v prípade dcérskeho uzla, ktorý nie je homogénne zaplnený konkrétnym symbolom, a teda tento uzol bude dekomponovaný a symbol 1 v prípade, ak je celý podstrom a teda aj jemu prislúchajúci subkvadrant homogénne zaplnený konkrétnym symbolom. Za týmto bitom 1 potom nasleduje binárna reprezentácia príslušného symbolu. Parameter n je nastavený na hodnotu 4 a parameter x je totožný s hodnotou COLSIZE.



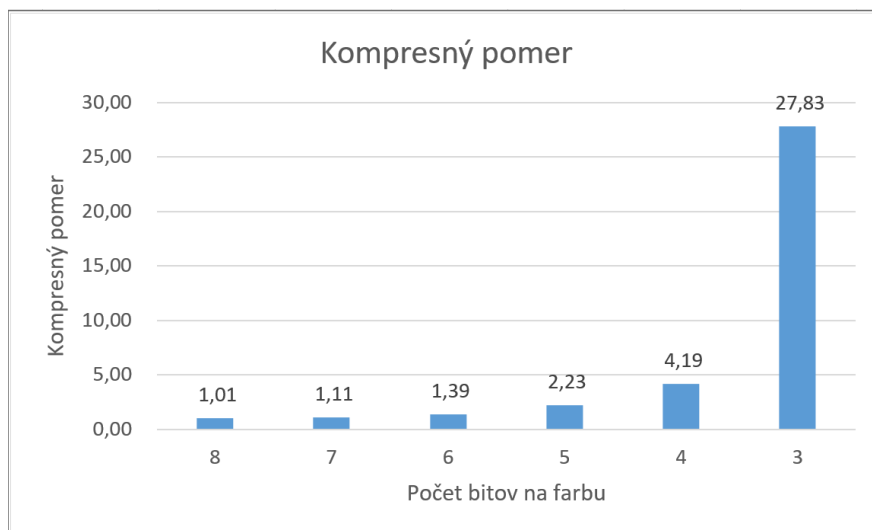
#### IV. VÝSLEDKY TESTOV

V rámci testovania grafického súborového formátu boli realizované testy na volumetrických datasetoch, z ktorých boli vybrané jednotlivé dvojdimenzionálne rezy a tie boli následne zakódované do navrhnutého grafického formátu s rôznym počtom odtieňov sivej.

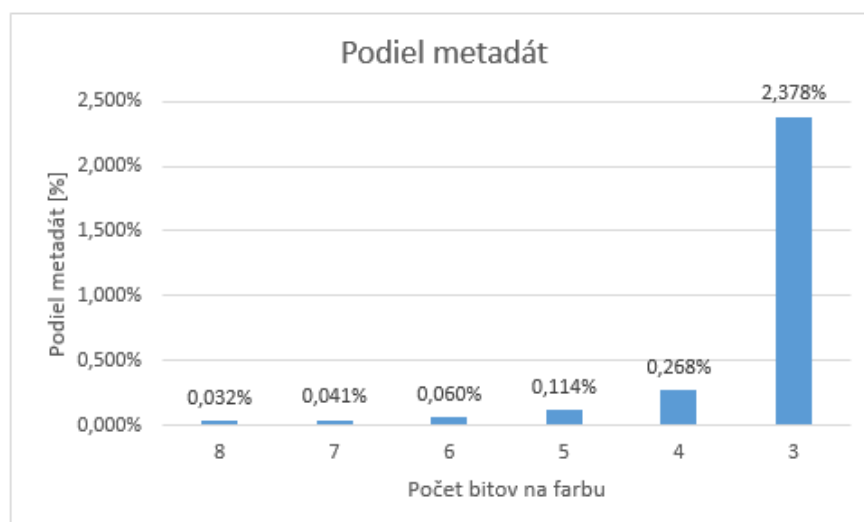
V prvom teste bol vybraný jeden rez z konkrétneho volumetrického datasetu a postupne bol znižovaný počet farieb z pôvodných 256 na 128, 64, 32, 16 a 8. Porovnávané potom bolo, koľko bitov zaberá tento obrázok zakódovaný v QIFF dátovom formáte a koľko bitov je v pôvodnej nekomprimovanej teda RAW podobe. Následne bol vypočítaný kompresný pomer a tiež bolo vyčíslené, koľko percent z QIFF formátu zaberá samotná hlavička súboru. Výsledky sú zhrnuté v Tabuľke 1 a Obrázku 1 a 2.

Tab. 1. Výsledky kompresie pri ukladaní dvojdimenzionálneho rezu testovacieho datasetu.

Počet farieb	b/farba	RAW	QIFF	Kompresný pomer	Podiel metadát
256	8	524 288	518 408	1,01	0,032%
128	7	458 752	412 528	1,11	0,041%
64	6	393 216	282 200	1,39	0,060%
32	5	327 680	146 912	2,23	0,114%
16	4	262 144	62 576	4,19	0,268%
8	3	196 608	7 064	27,83	2,378%



Obr. 1 Výsledky kompresie pri ukladaní dvojdimenzionálneho rezu testovacieho datasetu pre rôzne počty bitov na odtieň sivej (a teda počtu odtieňov).

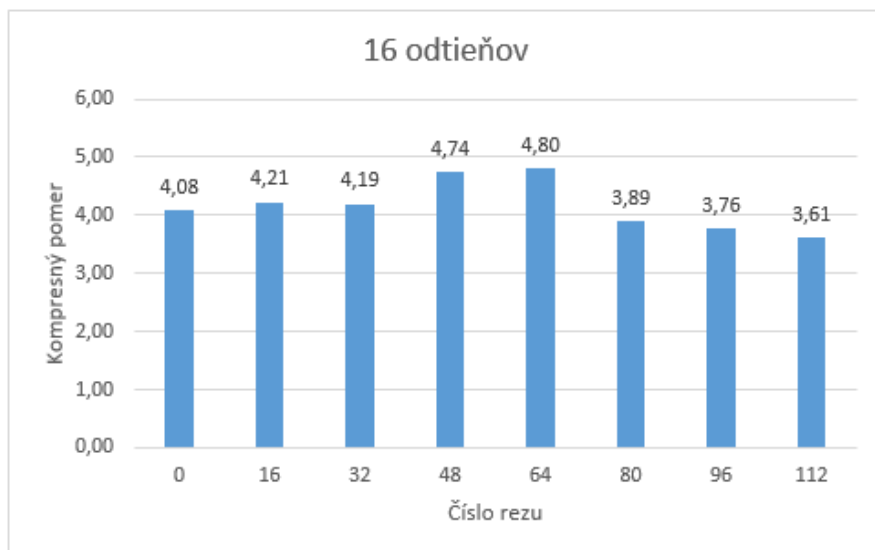


Obr. 2 Podiel metadát na celkovom objeme dát pri ukladaní dvojdimenzionálneho rezu testovacieho datasetu pre rôzne počty bitov na odtieň sivej (a teda počtu odtieňov).

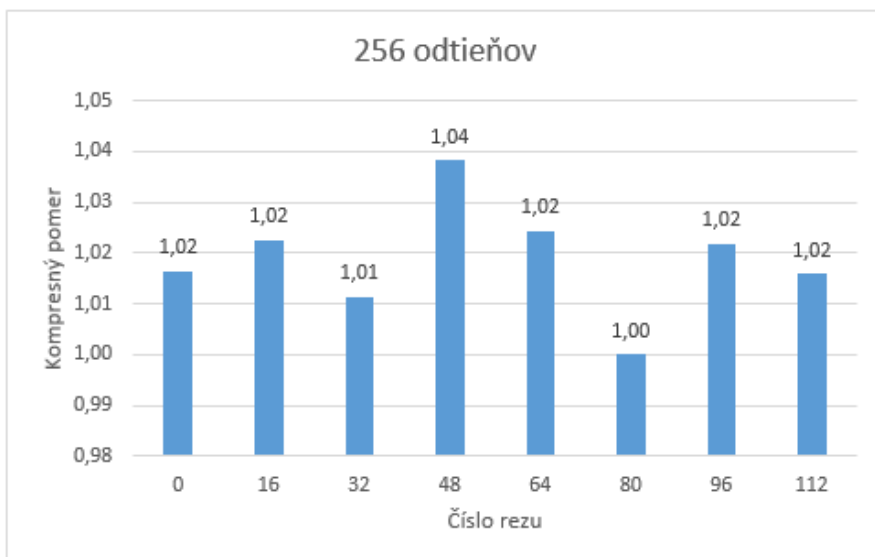
Ďalší test bol realizovaný na vybraných rezoch konkrétneho volumetrického datasetu, kde boli porovnávané výsledky kompresie pri ukladaní obrazu do formátu QIFF v porovnaní so surovými dátami RAW pre počet odtieňov sivej 256 a 16. Výsledky sú zhrnuté v Tabuľke 2 a na Obrázku 3 a 4.

Tab. 2 Výsledky kompresie pre viaceré rezy konkrétnym volumetrickým datasetom, pre 256 odtieňov sivej a 16 odtieňov sivej.

Rez	256 farieb			16 farieb		
	RAW	QIFF	Kompresný pomer	RAW	QIFF	Kompresný pomer
0	524 288	512 832	1,02	262 144	64 224	4,08
16	524 288	512 720	1,02	262 144	62 240	4,21
32	524 288	518 408	1,01	262 144	62 576	4,19
48	524 288	504 992	1,04	262 144	55 328	4,74
64	524 288	511 768	1,02	262 144	54 640	4,80
80	524 288	524 288	1,00	262 144	67 376	3,89
96	524 288	513 008	1,02	262 144	69 632	3,76
112	524 288	516 056	1,02	262 144	72 560	3,61



Obr. 3 Výsledky kompresie pre viaceré rezy konkrétnym volumetrickým datasetom, pre 16 odtieňov sivej.



Obr. 4 Výsledky kompresie pre viaceré rezy konkrétnym volumetrickým datasetom, pre 256 odtieňov sivej.

## V. ZÁVER

Cieľom tejto práce bolo navrhnúť iniciálnu verziu grafického súborového formátu pre ukladanie dvoj-dimenzionálnych dát s využitím kompresie dát pomocou hierarchických dátových štruktúr. Navrhnutý dátový formát umožňuje ukladať dvoj-dimenzionálne dáta, pričom každý obrazový bod má ako zaznamenaný atribút svoju farbu. Tá môže byť zapísaná priamo do dát obrázka, alebo je používaná paleta farieb. Ako hierarchická dátová štruktúra bola využitá dátová štruktúra založená na PSVO.

Výsledky testov ukázali, že je možné s využitím navrhnutého súborového formátu ukladať obrazovú informáciu a je možné dosahovať aj významný kompresný pomer. Ten je nízky v prípade, keď stúpa počet farieb, ktoré obsahuje farebná paleta, pretože to znižuje šancu na vytvorenie homogénnych plôch zložených z jednej farby. Naopak, tak ako klesá počet farieb vo farebnej palette, stúpa pravdepodobnosť existencie plôch zaplnených homogénne rovnakou farbou resp. odtieňom sivej. Ako ukazujú výsledky testov, v prípade použitia 16-tich odtieňov sivej bolo možné dosiahnuť kompresný pomer od 3,61 až po 4,80. V prípade použitia ôsmich farieb resp. odtieňov sivej bol maximálny dosiahnutý kompresný pomer až na úrovni 27,83.

Navrhnutý grafický formát je v súlade s predpokladmi vhodný pre účely uloženia obrázkov s obmedzeným počtom farieb a predpokladom existencie veľkých plôch homogénnej farby.

## POĎAKOVANIE

Tento výskum bol podporený projektom 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market.

## REFERENCIE

- [1] I. Gargantini, "An Effective Way to Represent Quadrees". In: Commun.ACM25.12 (dec. 1982), s. 905–910. ISSN: 0001-0782.doi:10.1145/358728.358741.
- [2] G. Hunter a K. Steiglitz, "Operations on Images Using QuadTrees". In: Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-1 (máj 1979), s. 145–153. doi:10.1109/TPAMI.1979.4766900.
- [3] A. Klinger a Ch. R. Dyer, "Experiments on picture representation using regular decomposition". In: Computer Graphics and Image Processing 5.1 (1976), s. 68–105. ISSN: 0146-664X. doi:https://doi.org/10.1016/S0146-664X(76)80006-8.
- [4] E. Kawaguchi a T. Endo, "On a Method of Binary-Picture Representation and Its Application to Data Compression". In: IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-2.1 (1980), s. 27–35. doi:10.1109/TPAMI.1980.4766967.
- [5] S. N. Srihari, "Representation of three-dimensional digital images". In: ACM Computing Surveys (CSUR)13.4 (1981), s. 399–424.
- [6] S. M. Rubin a T. Whitted, "A 3-Dimensional Representation for Fast Rendering of Complex Scenes". In: Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '80. Seattle, Washington, USA: Association for Computing Machinery, 1980, s. 110–116. ISBN: 0897910214. doi:10.1145/800250.807479. url:https://doi.org/10.1145/800250.807479.
- [7] Ch. L. Jackins a S. L. Tanimoto, "Octrees and their use in representing three-dimensional objects". In: Computer Graphics and Image Processing 14.370 (1980), s. 249–270. ISSN: 0146-664X. doi:https://doi.org/10.1016/0146-664X(80)90055-6.
- [8] R. E. Webber a M. B. Dillencourt, "Compressing quadtrees via common subtree merging". In: Pattern recognition letters 9.3 (1989), s. 193–200.
- [9] H. Samet, "Implementing ray tracing with octrees and neighbor finding". In: Computers & Graphics 13.4 (1989), s. 445–460. ISSN: 0097-8493. doi:https://doi.org/10.1016/0097-8493(89)90006-X.
- [10] D. Meagher, "Octree Encoding: A New Technique for the Representation, Manipulation and Display of Arbitrary 3-D Objects by Computer". In: (okt. 1980).
- [11] D. Meagher, "Geometric modeling using octree encoding". In: Computer graphics and image processing 19.2 (1982), s. 129–147.
- [12] D. Meagher, "The Octree Encoding Method for Efficient Solid Modeling". In: (aug. 1982).
- [13] S. Laine a T. Karras, "Efficient Sparse Voxel Octrees". In: Proceedings of the 2010 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. I3D '10. Washington, D.C.: Association for Computing Machinery, 2010, s. 55–63. ISBN: 9781605589398. doi:10.1145/1730804.17308.
- [14] V. Kämpe, E. Sintorn a U. Assarsson, "High Resolution Sparse Voxel DAGs". In: ACM Trans. Graph. 32.4 (júl 2013). ISSN: 0730-0301. doi:10.1145/2461912.2462024.71.
- [15] A. J. Villanueva, F. Marton a E. Gobbetti, "SSVDAGs Symmetry-Aware Sparse Voxel DAGs". In: Proceedings of the 20th ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. I3D '16. Redmond, Washington: Association for Computing Machinery, 2016, s. 7–14. ISBN: 9781450340434. doi:10.1145/2856400.285642.
- [16] L. Vokorokos, B. Madoš a Z. Bilanová, "PSVDAG: Compact Voxelized Representation of 3D Scenes Using Pointerless Sparse Voxel Directed Acyclic Graphs". In: Computing and Informatics (jan. 2020), s. 587–616. doi:10.31577/cai\_2020\_3\_587.
- [17] B. Madoš, N. Adam a M. Štancel, "Representation of Dense Volume Datasets Using Pointerless Sparse Voxel Octrees With Variable and Fixed-Length Encoding", 2021. In: 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI): 21. - 23.1.2021: Herľany, Slovakia, s. 343-348, Košice, Slovakia: Institute of Electrical and Electronics Engineers, 2021 ISBN 978-1-7281-8054-0.

# Categorical model of simply-typed lambda-calculus with reference type

<sup>1</sup>Tara STEFÁNYI, <sup>2</sup>Valerie NOVITZKÁ

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

<sup>1</sup>tara.stefanyi@student.tuke.sk, <sup>2</sup>valerie.novitzka@tuke.sk

**Abstract** – Categories are mathematical structures that primarily describe relationships between objects and not their very internal structure. We can use them to formalize problems from different areas, e.g. computer science or physics, into mathematics and solve the problems in another way. In this article, we will focus on the problem of reference types and work with dynamic memory, which is an integral part of programming languages. The aim is to create and evaluate a categorical model for a simple functional language extended by reference types.

**Keywords** – category theory, lambda calculus, reference type

## I. INTRODUCTION

The theory of categories uniformly describes structural concepts from various mathematical and computer science branches. It is not dealing with the internal structure of objects, but primarily with the relationships between the given objects. It is a powerful tool for mathematics and theoretical computer science, used to create more abstract structures. It can be considered as a universal language with which we study the properties of mathematical objects, try to understand the structure of mathematical theories and easily transfer ideas between different areas of mathematics.

Different areas have common patterns, trends or structures. When we have a problem in one area (such as topology) and we do not have the right tools to solve it, we can use categories that allow us to transfer the problem to another area (such as algebra) and see it in a different light. This way we can discover new tools to solve the problem and simplify the solution. Categories use abstraction to define and prove many complex and fine mathematical problems in a much simpler way [1]. The simplest example of a category is a set category, where objects are sets and morphisms are functions from one set to another.

We will take a closer look at the problem of reference types and work with dynamic memory, which is an integral part of creating programs, either the programmers have to manage memory themselves or the environment does it automatically. Dynamic memory (*heap*) is a type of memory where programs can store data during their execution, that is accessible even after the frame is replaced in local memory (*stack*), which occurs when calling or returning from a subprogram.

The aim of this work is to create a categorical model for reference types in lambda calculus and work with them. In the introductory part of the work we will focus on the basic theoretical foundations needed to create this categorical model.

## II. REFERENCE TYPES

In this section, we provide basic information, properties, and operations that we can perform with reference based on [2], [3].

*Reference types* are types that have side effects. References describe particular cells in memory, the address we want to access. By reference types we mean a set of references also with operations on them. The advantage of reference types is the ability to work with structures that have different dynamic lengths. One example of the use of references are link lists. For the formal definition of reference types, it is necessary to introduce memory abstraction, as reference types work with dynamic memory [4].

The following basic operations are defined on the values of the reference types:

- *allocation* to reserve a memory cell, using the operator `alloc`,

- *dereference* to obtain the value stored at the address where the reference points, using the operator @,
- *assignment*, thus, modifying the contents of the cell referenced by the operator :=.

### Reference type

- allows working with dynamic memory,
- has basic operations (allocation `alloc`, dereference @, assignment :=),
- is called *Ref T*, where *T* is the type of value of the cell we are referring to,
- using dereference we get the contents of the cell we are referring to,
- by assignment we store the value in the referenced cell, the result of the assignment is of type `Unit`.

We will rely on the garbage collector to free memory because it is very challenging to achieve type security in an explicit deallocation operation.

#### A. Memory abstraction

In the formal definition of the language used in this work, we will work with an abstraction of dynamic memory. One view of memory is that it consists of a large array of bytes and the operating system remembers which parts of the array are already used. The representation of types might be different in different architectures. In order not to have to deal with that we will modify the view of memory when formally defining a language:

- dynamic memory will be an array of positions instead of an array of bytes,
- the reference will be an element of uninterpreted type at a position in memory,
- dynamic memory will be a partially defined function that assigns a value stored at that position to a memory position.

### III. BASIC CONCEPTS OF CATEGORY THEORY

This section will mention some basic category theory used in this paper introduced in many sources such as [5], [6], [7], [8]. We will define a category as a mathematical structure for which there are several equivalent definitions. We will list one of them based on [5], [7], since it does not require graph theory, for example in [7], [9].

*Definition 1:* A category  $\mathcal{C} = (\mathcal{C}_{obj}, \mathcal{C}_{morp})$  is a structure, which consists of:

- class of objects  $\mathcal{C}_{obj}$ ,
- class of morphisms  $\mathcal{C}_{morp}$ ,

where each morphism has a specified domain *dom* and a codomain *cod* where notation  $f : A \rightarrow B$  says that *f* is a morphism from the domain *A* to the codomain *B*.

The following conditions must apply in a category [7]:

- If  $f : A \rightarrow B$  is a morphism in a category  $\mathcal{C}$ , then  $dom(f) = A$  and  $cod(f) = B$ , so that  $dom(f)$  is the domain and  $cod(f)$  is the codomain of morphism *f*.
- For each object *A* there is an identical morphism  $id_A : A \rightarrow A$ .
- If  $f : A \rightarrow B$  and  $g : B \rightarrow C$  are morphisms then there is morphism  $f \circ g : A \rightarrow C$  which is a composition of morphisms *f* and *g*.
- The composition of morphisms is associative, e.g. for morphisms  $f : A \rightarrow B$ ,  $g : B \rightarrow C$ ,  $h : C \rightarrow D$  applies  $f \circ (g \circ h) = (f \circ g) \circ h$ .

#### A. Functors

*Definition 2:* Functor  $F : \mathcal{C} \rightarrow \mathcal{D}$  is a pair of functions  $(F_0, F_1)$

$$\begin{aligned} F_0 : \mathcal{C}_{obj} &\rightarrow \mathcal{D}_{obj} \\ F_1 : \mathcal{C}_{morp} &\rightarrow \mathcal{D}_{morp} \end{aligned}$$

for which the following conditions apply:

- If  $f : A \rightarrow B$  is a morphism in  $\mathcal{C}$  then  $F_1(f) : F_0(A) \rightarrow F_0(B)$  in  $\mathcal{D}$ .
- For any object *A* in  $\mathcal{C}$  applies  $F_1(id_A) = id_{F_0(A)}$ .
- If composition  $f \circ g$  is in  $\mathcal{C}$ , then composition  $F_1(f) \circ F_1(g)$  is defined in  $\mathcal{D}$  and  $F_1(f \circ g) = F_1(f) \circ F_1(g)$  applies.

Functors preserve morphisms, identities and compositions of morphisms.

*Definition 3:* Functor  $F : \mathcal{C} \rightarrow \mathcal{C}$  is called an *endofunctor* on a category  $\mathcal{C}$ . In addition, if for any object *C* and morphism *f* in  $\mathcal{C}$

$$\begin{aligned} FC &= C \\ Ff &= f, \end{aligned}$$

applies, then *F* is an identical functor on  $\mathcal{C}$ .

## IV. A CATEGORICAL MODEL OF REFERENCE TYPES

In this section we will create a categorical model of reference types with separated static and dynamic memory. There will be a separate category for the dynamic memory where the allocation and dereference operations will be functors and the assignment operation will be an endofunctor over the dynamic memory.

We will build on the category model of typed  $\lambda$ -calculus with the syntax as follows:

$$\begin{aligned}
t ::= & x \mid \lambda x : T.t \mid tt \mid \text{if } t \text{ then } t \text{ else } t \mid \text{succ } t \mid \text{pred } t \mid \text{iszero } t \mid v \\
& \mid \langle t, t \rangle \mid t.1 \mid t.2 \mid \text{unit} \\
v ::= & \lambda x : T.t \mid \text{true} \mid \text{false} \mid nv \mid \langle v, v \rangle \mid \text{unit} \\
nv ::= & 0 \mid nv \\
T ::= & \text{Bool} \mid \text{Nat} \mid T \rightarrow T \mid T \times T \mid \text{Unit}
\end{aligned} \tag{1}$$

For the formal definition of this category see [10] or [11]. The main idea is that the objects of the category correspond to types and morphisms to operations of lambda calculus. In this paper we will call this category  $\mathcal{C}$ .

## A. Principles, motivation, basic ideas

Our category model will be made of two separate categories. First one is the category model  $\mathcal{C}$  of typed  $\lambda$ -calculus mentioned above. In this paper we will only list the part of the model that relates to reference types, and thus the following syntax [11]:

$$\begin{aligned}
t ::= & \dots \mid \text{alloc } t \mid @t \mid t := t \mid l \\
v ::= & \dots \mid l \\
T ::= & \dots \mid \text{Ref } T
\end{aligned} \tag{2}$$

which is the extension of the simple typed lambda calculus with reference types. We will add reference types by creating a second model  $\mathcal{D}$  that will represent the dynamic memory and the added operations will be functors and endofunctors.

As we have to work with the dynamic memory, we model the abstraction of computer memory as a table of memory cells. Memory cells will be indexed by positions and values will be stored in cells. First of all, we need to define positions as a set of positions:

$$\text{Loc} = \mathbb{N} \tag{3}$$

where elements  $l \in \text{Loc}$  are positions in the dynamic memory.

For each type object, we will also have an object for the relevant reference type, which will also carry the type information. The elements of dynamic memory will be simple as the values of different types stored in dynamic memory will be in different objects. In the model, we will have in theory an infinite number of such objects, while dynamic memory is their unification. The representation of reference types in the category will look as follows:

$$\begin{aligned}
\text{Dynamic } \mathbb{N} &= \text{Loc} \times \mathbb{N}, \\
\text{Dynamic } \mathbb{B} &= \text{Loc} \times \mathbb{B}, \\
\text{Dynamic } T_1 &= \text{Loc} \times \llbracket T_1 \rrbracket, \\
\text{Dynamic } T_2 &= \text{Loc} \times \llbracket T_2 \rrbracket, \\
&\vdots \\
\text{Dynamic } T_n &= \text{Loc} \times \llbracket T_n \rrbracket.
\end{aligned} \tag{4}$$

The elements of reference type will contain the position in memory and the value of the type that is stored in the given position. For any type representation  $\llbracket T_i \rrbracket$  the elements will be:

$$(l_i, v_i), \tag{5}$$

where  $v_i \in \llbracket T_i \rrbracket$ . The dynamic memory for this type is:

$$\text{Dynamic } T_i = \langle (l_{i1}, v_{i1}), \dots, (l_{imax}, v_{in}) \rangle, \tag{6}$$

where  $l_i \in \text{Loc}$  is the position or address of a given memory cell for type  $\llbracket T_i \rrbracket$  and  $v_i$  is a value of type  $\llbracket T_i \rrbracket$ , which is a type from a set of type representations  $\llbracket Types \rrbracket$ :

$$\llbracket Types \rrbracket = \{ \llbracket T_1 \rrbracket, \dots, \llbracket T_n \rrbracket \}. \tag{7}$$

Dynamic memory will be defined as unification of objects  $\text{Dynamic } T$  for each type which has a set of positions and values at a given position. In theory the number of reference objects

would be infinite, as every type can be constructed of other types. For  $l_{ij} \in \text{Loc}$  and  $\llbracket t \rrbracket \in \llbracket T_i \rrbracket$  and the actual memory  $dyn_1 \in \text{Dynamic } T_1$ ,  $dyn_i \in \text{Dynamic } T_i$  to  $dyn_s \in \text{Dynamic } T_s$  will be defined by the following regulation:

$$\begin{aligned} dyn_1 &= \langle (l_{11}, v_{11}), \dots, (l_{1j}, v_{1j}), \dots, (l_{1n}, v_{1n}) \rangle \\ &\vdots \\ dyn_i &= \langle (l_{i1}, v_{i1}), \dots, (l_{ij}, v_{ij}), \dots, (l_{in}, v_{in}) \rangle \\ &\vdots \\ dyn_s &= \langle (l_{s1}, v_{s1}), \dots, (l_{sj}, v_{sj}), \dots, (l_{sn}, v_{sn}) \rangle \end{aligned} \quad (8)$$

### B. Construction of the categorical model

As we mentioned before the model will consist of two categories  $\mathcal{C}$  and  $\mathcal{D}$ . First is the category  $\mathcal{C}$  representing types lambda calculus with the following objects:

$$\mathbb{N}, \mathbb{B}, \llbracket T_1 \rrbracket, \dots, \llbracket T_n \rrbracket, \text{Loc} \quad (9)$$

and the morphisms correspond to the functions of the typed lambda calculus. We will not give a definition of this category, as this is not the aim of this work.

Now we can construct the second category  $\mathcal{D}$  representing reference types. We will define its objects and morphisms.

### C. Objects

Category  $\mathcal{D}$  will have in theory an infinite number of objects of dynamic memory:

$$\begin{aligned} &\text{Dynamic } \mathbb{N}, \\ &\text{Dynamic } \mathbb{B}, \\ &\text{Dynamic } T_1, \\ &\vdots \\ &\text{Dynamic } T_n. \end{aligned} \quad (10)$$

### D. Functors

We introduce allocation, dereference and assignment operations as functors and endofunctors. For allocation we define a functor  $A : \mathcal{C} \rightarrow \mathcal{D}$

$$A_0 = \mathcal{C}_{obj} \rightarrow \mathcal{D}_{obj} \quad (11)$$

The allocation functor works on objects as follows:

$$A_0 \llbracket T \rrbracket \llbracket t \rrbracket = \begin{cases} \langle (l_{i1}, v_{i1}), \dots, (l_{in}, v_{in}), (l_{i(n+1)}, \llbracket t_1 \rrbracket) \rangle & \text{if } \llbracket t \rrbracket \in \llbracket T \rrbracket \text{ a } t = \text{alloc}(\llbracket t_1 \rrbracket) \\ \perp & \text{otherwise.} \end{cases}$$

$$A_1 = \mathcal{C}_{morp} \rightarrow \mathcal{D}_{morp} \quad (12)$$

$$A_1 \llbracket m \rrbracket = \begin{cases} \text{Dynamic } T \rightarrow \perp & \text{if } m = \llbracket T \rrbracket \rightarrow \llbracket T' \rrbracket \\ id_{\perp} & \text{otherwise.} \end{cases}$$

We will also create a functor that inserts the occupied value into the position object and the one that returns the position occupied by the function `alloc`.

$$L_0 = \mathcal{D}_{obj} \rightarrow \mathcal{C}_{obj} \quad (13)$$

The functor for expressing the last occupied position works on objects as follows:

$$L_0 \llbracket T \rrbracket \llbracket t \rrbracket = \begin{cases} l_{in} & \text{if } dyn_i \in \text{Dynamic } T \text{ a } t = \text{lastPos}(dyn_i) \\ \perp & \text{otherwise.} \end{cases}$$

$$L_1 = \mathcal{D}_{morp} \rightarrow \mathcal{C}_{morp} \quad (14)$$

$$L_1 \llbracket m \rrbracket = \begin{cases} \text{Dynamic } T \rightarrow \perp & \text{if } m = \llbracket T \rrbracket \rightarrow \llbracket T' \rrbracket \\ id_{\perp} & \text{otherwise.} \end{cases}$$

For dereference we define a functor  $D : \mathcal{D} \rightarrow \mathcal{C}$

$$D_0 = \mathcal{D}_{obj} \rightarrow \mathcal{C}_{obj} \quad (15)$$

The dereference functor works on objects as follows:

$$D_0[[T]][t] = \begin{cases} v_{ij} & \text{if } t = @([l_{ij}]) \\ \perp & \text{otherwise.} \end{cases}$$

$$D_1 = \mathcal{D}_{morp} \rightarrow \mathcal{E}_{morp} \quad (16)$$

$$D_1[[m]] = \begin{cases} \text{Dynamic } T \rightarrow \perp & \text{if } m = [T] \rightarrow [T'] \\ id_{\perp} & \text{otherwise.} \end{cases}$$

We define an endofunctor for the assignment  $H : \mathcal{D} \rightarrow \mathcal{D}$

$$H_0 = \mathcal{D}_{obj} \rightarrow \mathcal{D}_{obj} \quad (17)$$

The assignment endofunctor works on objects as follows:

$$H_0[[T]][t] = \begin{cases} \langle (l_{i1}, v_{i1}), \dots, (l_{ij}, [t_1]), \dots, (l_{in}, v_{in}) \rangle & \text{if } t = [l_{ij} := t_1] \\ id_{\perp} & \text{otherwise.} \end{cases}$$

$$H_1 = \mathcal{D}_{morp} \rightarrow \mathcal{D}_{morp}. \quad (18)$$

$$H_1[[m]] = \begin{cases} id_{\text{Dynamic } T} & \text{if } m = [T] \rightarrow [T'] \\ id_{\perp} & \text{otherwise.} \end{cases}$$

The allocation, dereference, and assignment functions will be defined using functors as follows:

$$[[\text{alloc}]]t = L \circ \text{lastPos}(A([[ \text{alloc} ]][t_1])) \quad (19)$$

$$[[@]]l_{ij} = D([[@]]l_{ij}) \quad (20)$$

$$[[l_{ij} := t]] = H([[l_{ij} := t_1]]) \quad (21)$$

The category model with functors, which we created in this section is shown in the figure (1).

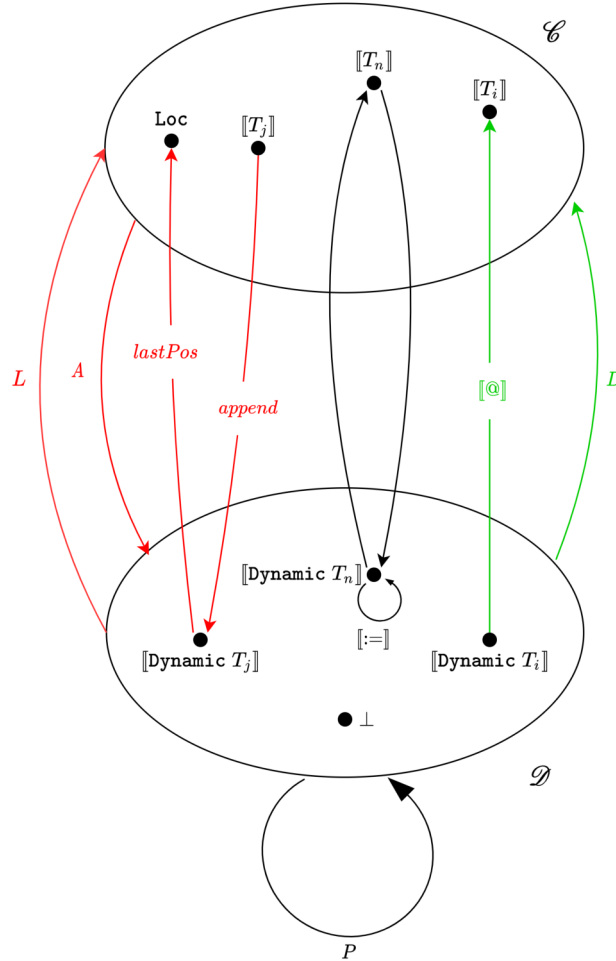


Fig. 1 Model with separated memory



## V. CONCLUSION

The aim of this work was to create a category and its model for a typed lambda calculus extended by a reference type. In this work we managed to create a categorical model where we have a separate object for each reference type. The number of objects in the category therefore depends on the number of types of values we use. When using a large number of different types we will have many separate objects for the dynamic memory. The advantage is that it is not necessary to store the type for each value. There is no problem with memory fragmentation in this model, because we can easily replace free cells, since each dynamic memory object has its own type. Another continuation of this work could be a further modification of this model. The types of lambda calculus could be written as monads and dynamic memory objects from the second category as well, making the category even more abstract. In this form, it would be easier to compare this category with other areas.

## ACKNOWLEDGMENT

This work was supported by the Project KEGA 011TUKE4/2020: “A development of the new semantic technologies in educating of young IT experts.”

## REFERENCES

- [1] R. Geroch, *Mathematical Physics*, ser. Chicago Lectures in Physics. University of Chicago Press, 1985.
- [2] P. Naumov, “Formalizing reference types in nuprl,” Ph.D. dissertation, Cornell University, 1998.
- [3] —, “Theory of reference types,” *Cornell University*, 1998.
- [4] D. Mihályi and V. Novitzká, *Teória typov*. Technical University of Košice, 2015.
- [5] A. Asperti and G. Longo, *Categories, Types and Structures*. M.I.T. Press, 1991.
- [6] S. Awodey, *Category Theory*, 1st ed. Clarendon Press ; Oxford University Press, 2006.
- [7] V. Novitzká and V. Slodičák, *Kategorické štruktúry a ich aplikácie v informatike*. Technical University of Košice, 2010.
- [8] D. Turi, “Category theory lecture notes,” *University of Edinburgh*, 2001.
- [9] M. Barr and C. Wells, *Category Theory for Computing Science*. USA: Prentice-Hall, Inc., 1990.
- [10] R. L. Crole, *Categories for Types*. Cambridge University Press, 1994.
- [11] V. Novitzká, W. Steingartner, and J. Perháč, “A simple categorical model of reference type,” *The IPSI Transactions on Advanced Research*, vol. 17, no. 1, 2021.

# Vytvorenie prototypu na spracovanie transakčných dát v MSP pomocou jazyka R

Matúš KRIZÁNEK, František BABIČ, Viera PAVLIŠKOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Letná 9, 040 01 Košice, Slovenská republika

matus.krizanek@student.tuke.sk, frantisek.babic@tuke.sk, viera.pavliskova@tuke.sk

**Abstrakt** — Čoraz viac firiem, ktoré využívajú prostriedky na podporu rozhodovania, spoznáva nákupné správanie zákazníkov a hľadá najvhodnejšie formy marketingových aktivít pomáhajúcich udržať firmu konkurencieschopnou. Do návrhu prototypu sme zahrnuli základný balík požiadaviek na tradičné riešenie Manažérskeho informačného systému (MIS) v podmienkach malej alebo strednej firmy (MSP): sumárny pohľad tržieb pre jednotlivé prevádzky, pohľad na nákupy zákazníkov, filtrovanie jednotlivých nákupov, hľadanie v nákupoch zákazníkov, najviac predávané produkty, štatistiky, tabuľkové výstupy pre jednotlivé prevádzky a pre celú firmu v sumárnej podobe. Na implementáciu sme použili voľne dostupné a moderné technológie. Úspešnosť testovania ponúka námet pre vytvorenie biznis plánu, ktorý by pokrýval dokončenie a rozšírenie produktu pre rôzne MSP na Slovenskom ale aj zahraničnom trhu.

**Kľúčové slová** — transakčné dáta, jazyk R, reporty

## I. ÚVOD

„Aké produkty kupujú naši zákazníci spoločne?“ je častou otázkou v podnikaní. Maloobchodníci vyžadujú na túto otázku odpoveď, pretože im to poskytuje poznatky o nákupnom správaní zákazníkov. Takéto poznatky vedú k strategickým zmenám v usporiadaní obchodov, výrobkov, ktoré sú na sklade a k marketingovým kampaniam. Na analýzu obchodných transakcií nakupujúcich zákazníkov používajú maloobchodníci rôzne prístupy štatistického modelovania [1]. Podniková inteligencia (BI) označuje technológie, aplikácie a postupy zhromažďovania, integrácie, analýzy a prezentácie obchodných informácií. Spoločnosti sú nútené zhromažďovať, chápať a využívať svoje údaje na podporu rozhodovania s cieľom zlepšiť obchodné operácie [2]

V našom článku sme sa zamerali na návrh vlastného prototypu umožňujúceho spracovanie transakčných dát v MSP prostredníctvom voľne dostupného jazyka R ako odpoveď na existujúce konkurenčné riešenia a identifikované kľúčové požiadavky používateľov.

### A. Prehľad Vybraných Existujúcich riešení

Pri porovnávaní existujúcich riešení sme sa zamerali na prípadové štúdie niektorých projektov, ktoré boli realizované v reálnom prostredí pomocou softvérových riešení Tableau, Microsoft Power BI a Qlik Sense. V jednotlivých štúdiách sme skúmali cieľ, požiadavky, návrh riešenia, nasadenie a vyhodnotenie projektu.

V roku 2016 začala spoločnosť Lufthansa riešiť možnosť centralizovanejšieho, štandardizovanejšieho a jednotného riešenia pre vykazovanie údajov. Spoločnosť v tom čase nedisponovala jednotným spôsobom vykazovania jednotlivých oddelení. Čas a vynaložené úsilie potrebné na prípravu reportov boli náročné; IT tímy neboli schopné orientovať dostatok pracovného času na analýzu a interpretáciu údajov. Lufthansa si vybrala nástroj Tableau. Spočiatku bolo toto riešenie nasadené v obchodnom oddelení a neskôr bolo súčasťou celej firmy. V súčasnosti vo firme využíva toto riešenie viac ako 500 ľudí, ktorí pracujú so správami z Tableau a analyzujú ich prostredníctvom Tableau servera a príslušných dashboardov [3]. Podľa autorov štúdie nasadenie tohto riešenia prinieslo zvýšenie efektívnosti spoločnosti o 30%, pretože sa výrazne znížil čas na prípravu údajov a generovanie správ.

V druhej prípadovej štúdii autor článku píše o spolupráci so zákazníkom, ktorý vyrába plastové výrobky [4]. Firma mala ťažkosti s monitorovaním vyťaženia pracovnej sily a jej cieľom bolo zlepšiť vykazovanie v tejto oblasti. V spolupráci so spoločnosťou Columbus navrhli riešenie využívajúce funkcie Power BI na príjem dát do štandardizovaného modelu. Microsoft Power BI poskytol servisnému tímu široký prehľad o ich zdrojoch, o tom, ako sú pridelené a aký to má vplyv na plánované úlohy. Táto správa sa aktualizuje osemkrát denne, aby sa zabezpečil prístup

k informáciám o využití zamestnancov v reálnom čase. V nástroji Power BI používajú podrobné dashboards a prispôsobiteľné správy, aby lepšie porozumeli údajom pre alokáciu prostriedkov.

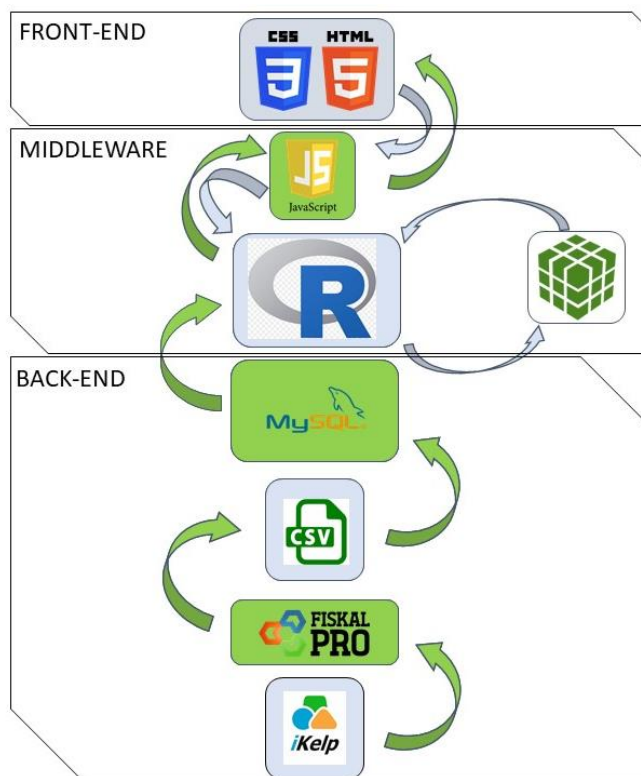
Firma Nexus prevádzkuje metro, trajektové služby a poskytuje podporu pre autobusovú dopravu a infraštruktúru [5]. Spoločnosť Perfectx bola prvýkrát oslovená spoločnosťou Nexus v roku 2010 s cieľom poskytnúť robustné riešenie BI. Nexus sa rozhodol pre kombináciu služieb Perfect Image a softvéru Qlik. Predtým oddelenie BI trávilo veľa času zhromažďovaním údajov o predaji lístkov a odpovedaním na základné žiadosti o informácie oddelenia predaja lístkov. Teraz je tento proces zabezpečený pomocou nástroja Qlik. Zároveň nové riešenie nahradilo nutnosť extrahovania údajov z mnohých rôznych systémov a ich manuálne čistenie, formátovanie a zlúčenie. Tento prístup pomohol aj pri riešení dlhodobého problému verejnosti s automatmi na lístky. Na základe získaných dát boli automaty aktualizované a prispôbené požiadavkám a očakávaniam zákazníkov.

## II. TVORBA VLASTNÉHO RIEŠENIA

Na spracovanie transakčných dát v podniku sme navrhli prototyp ktorý zastrešuje potrebnú funkcionality stanovenú používateľskými požiadavkami. Navrhnutý prototyp spracováva a následne vizualizuje zozbierané dáta do interaktívnej podoby vo forme používateľských obrazoviek. Prototyp by mal poslúžiť ku komplexnému pohľadu na dáta z jednotlivých firemných prevádzok, ktoré doposiaľ vo firme neboli spracované. Spoznanie svojich zákazníkov dodáva firme pridanú hodnotu, vďaka ktorej sa stáva konkurencieschopná veľkým obchodným reťazcom.

### A. Architektúra

Architektúra prototypu je rozdelená na 3 úrovne:



Obr. 1 Architektúra navrhnutého riešenia

### B. Implementácia

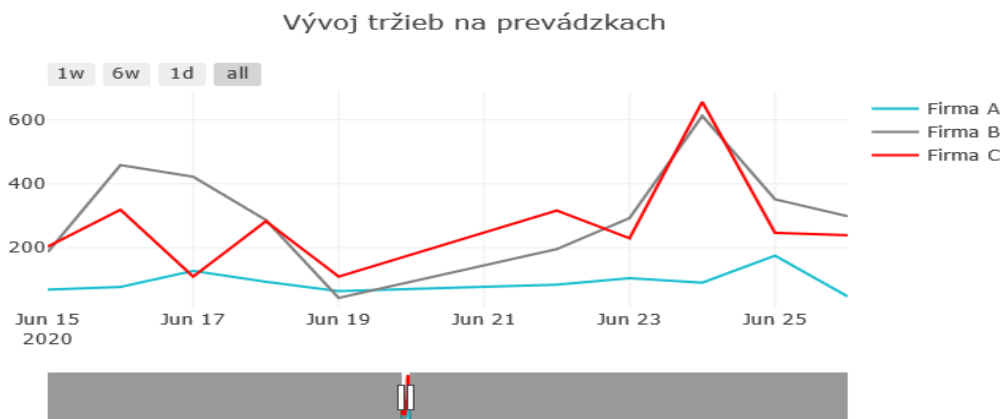
Zozbierané dáta z pokladničných systémov sú posielané na cloudovú službu, ktorá poskytuje prístup k pokladničným dokladom v reálnom čase. Tieto dáta zaznamenávajú aj informácie o zákazníkovi číslo každého evidovaného zákazníka pridelené k jednotlivým položkám nákupov. Tieto informácie o nákupoch sme transformovali do lokálnej databázy, ku ktorej sme neskôr pristupovali v programovacom prostredí RStudio jazyka R. Pre prácu s jazykom R sme si stiahli knižnicu RMariaDB, ktorá umožňuje prístup k databáze; data.table pre prácu s dátovými tabuľkami ako pomocnými tabuľkami; data.cube pre vytvorenie dátovej kocky a ďalšie pomocné knižnice. Vytvoreniu dátovej kocky predchádzala séria krokov, ktorú bolo potrebné vykonať prv, aby bolo možné následne s dátovou kockou pracovať. Prvým krokom bolo namodelovanie toho, ako výsledná kocka bude vyzeráť.

Definovali sme si jednotlivé hierarchie ako prípravu pre vytvorenie dátovej kocky. Prvú hierarchiu predstavuje dátum, ktorý sa člení na rok, mesiac, týždeň v roku a názov dňa v týždni. Druhú hierarchiu predstavuje firma, ktorá sa ďalej člení na jednotlivé prevádzky. Treťou hierarchiou je názov tovarov, ktoré patria do určitých skupín produktov. Poslednou zloženou hierarchiou dátovej kocky je doklad. Ten sa člení z hľadiska čísla zákazníka priradenému k číslu uzávierky, položky s definovaným množstvom zahŕňajúcim jednotkovú cenu produktu a číslom dokladu s priradeným typom platby.

Pri vytváraní hierarchií bol problém s nedostatočnou dokumentáciou v použitej knižnici, pretože príklady a ukážky boli dostupné iba k jednoduchým dátam bez použitia zloženej hierarchie. Ďalšie príklady použitia neboli dostupné a preto bolo potrebné experimentovať s algoritmom použitým v danej knižnici.

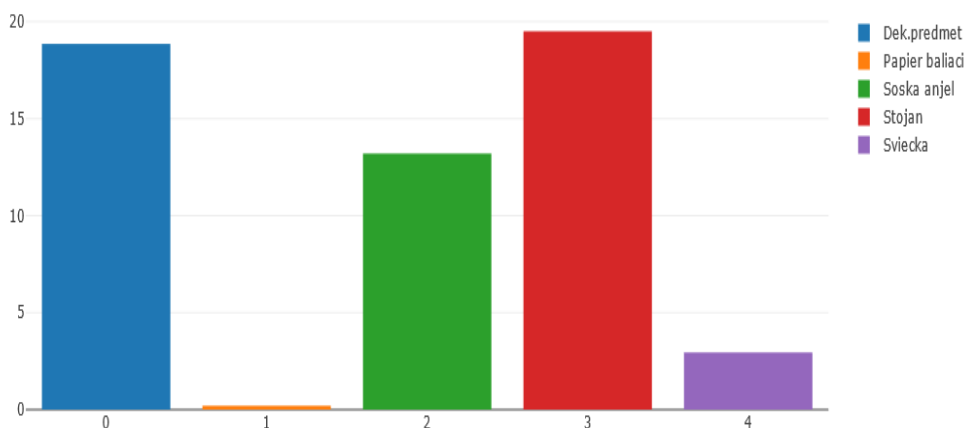
Prostredníctvom Rest API vytvoreného v jazyku R sme definovali presmerovania a GET funkcie pre prístup front-endu k back-endu prototypu. Middleware tvorí programovací jazyk JavaScript, ktorý slúži pre prihlasovanie do aplikácie a pomocou knižnice Plotly vykresľujeme interaktívne prostredie pre front-end. Prihlasovanie do aplikácie je zabezpečené pomocou služby Firebase od spoločnosti Google.

Front-end aplikácie využíva bootstrap, ktorého kaskádové štýly sú vkladané pomocou CDN uložených v cloudovej službe sprístupnenej odkazmi. To isté platí pre niektoré z funkcií JavaScriptu. To znamená, že čistý HTML súbor stránky nepotrebuje k svojej funkcionalite žiadne interne uložené súbory, na ktoré je potrebné prepojenie. Prototyp pracuje ako responzívna webová aplikácia. Vytvorený prototyp je zložený z navigácie, v ktorej je možné kliknúť na Dashboard, ktorý predstavuje sumárny pohľad na všetky prevádzky, a ďalšie záložky obsahujúce pohľady na jednotlivé prevádzky. Na obrázku (Fig. 2) je zobrazený graf porovnania vývoja tržieb na jednotlivých prevádzkach za zvolené obdobie.



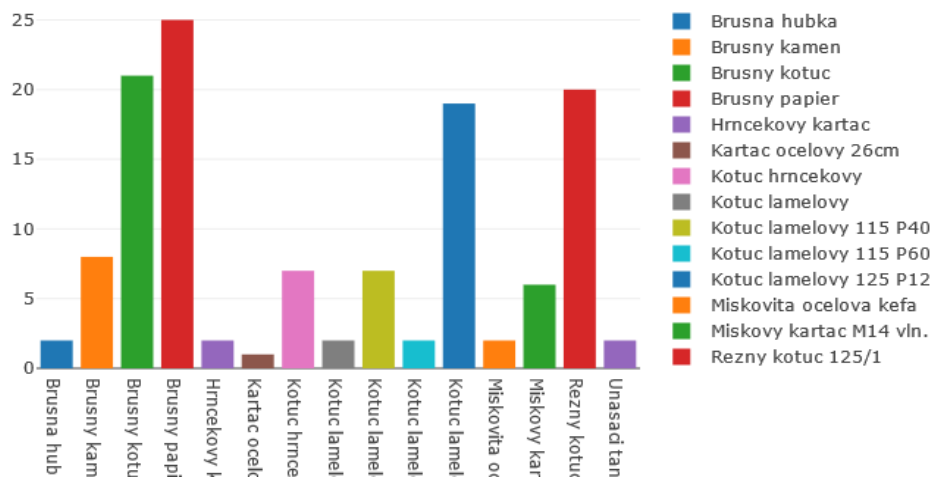
Obr. 2 Porovnanie tržieb na prevádzkach

Ďalšiu funkcionálnosť predstavuje porovnanie najviac predávaných produktov z vybranej skupiny tovarov za zvolené časové obdobie (Fig. 3).



Obr. 3 Najviac predávané produkty

Na zobrazenie zákaznických nákupov slúži ďalšia funkcia, pomocou ktorej používateľ môže sledovať nákupné správanie zákazníka na prevádzke, alebo v celej firme (Fig. 4).



Obr. 4 Nákupy zákazníka

### C. Testovanie

Vytvorený prototyp sme otestovali simuláciou operačného prostredia vo vybranej firme. Zorganizovali sme testovanie so šiestimi potenciálnymi používateľmi, ktorí zastupovali vedenie, operatívny manažment a vedúcich jednotlivých prevádzok. Pri testovaní sme sa zamerali na oboznámenie sa s prototypom a vypracovanie vopred definovaných úloh. Z pozorovania a vedenia testovania sme zaznamenávali, s čím mali používatelia problém, čo je potrebné pridať alebo odstrániť z aplikácie, aby bola používateľsky prívetivá. Na testovanie prototypu bol vyčlenený čas 6 hodín. Každý používateľ mal k dispozícii 60 minút pre vypracovanie jednotlivých úloh, individuálny rozhovor a vyplnenie dotazníka spokojnosti.

Vedenie firmy a zamestnanci zhodnotili, že takými znalosťami zo zákaznických nákupov nikdy pred tým pri rozhodovaní nedisponovali. Prototyp hodnotili ako veľmi prospešný a prínosný pre ďalšie fungovanie firmy. Vďaka využitým open-source technológiám bežne používaným v reálnej praxi a najnovším verziám knižníc s technickou podporou má prototyp dobrý potenciál pre bezproblémové nasadenie do rôznych MSP.

### III. ZHRNUTIE

Na trhu je dostupných množstvo softvérových nástrojov, ktoré riešia podporu v rozhodovaní pre prostredie firiem, avšak nie každé riešenie vyhovuje potrebám a predstavám všetkých podnikov. Nasadenie vytvoreného prototypu vo firme poskytuje prehľad o transakciách, zákaznických nákupoch a dodáva vhodné výstupy, ktoré môžu byť použité pre ďalšie rozhodovanie.

Implementácii navrhnutého prototypu predchádzal čas strávený analýzou, návrhom, tvorbou a testovaním produktu v časovom ponímaní približne 160 hodín. Z hľadiska nasadenia produktu by predaj riešenia na mieru pre firmu pozostával z: analýza potrieb zákazníka (200€), spracovanie transakčných dát (800€) a predaj licencie na produkt + zaškolenie personálu (1000€). Každoročný prenájom obnovenej licencie a hostingu by predstavoval sumu 250€/ 1 rok. Úspešnosť testovania ponúka námet pre vytvorenie biznis plánu, ktorý by pokrýval dokončenie a rozšírenie produktu pre rôzne MSP na Slovenskom ale aj zahraničnom trhu.

### ZDROJE

- [1] M.A Adnan, "A set-based visual analytics approach to analyze retail data". EuroVA '18: Proceedings of the EuroVis Workshop on Visual Analytics, 2018, 37-41.
- [2] L. Hurbean, "Business intelligence: applications, trends, and strategies", Analele Stiintifice ale Universitatii "Alexandru Ioan Cuza" din Iasi - Stiinte Economice (1954-2015), 52-53, 2006, 307-312.
- [3] Top Tableau Case Studies – JP Morgan, Lenovo & Lufthansa, 2021, <https://data-flair.training/blogs/tableau-case-studies/>.
- [4] Power BI in the manufacturing sector: A real-life case study, 2021, <https://www.columbusglobal.com/en-gb/blog/power-bi-in-the-manufacturing-sector-a-real-life-case-study>.
- [5] Nexus + Perfect Image: a case study: <https://perfect-image.co.uk/nexus/>

# Evidencia dočasného parkovania pomocou metód počítačového videnia a umelej inteligencie

*Kristián MIČKO, František BABIČ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Letná 9, 040 01 Košice, Slovenská republika

kristian.micko@student.tuke.sk, frantisek.babic@tuke.sk

**Abstrakt** — Správa parkovacích miest je v súčasnej dobe veľmi rozšírený problém najmä v mestách. Tento problém je v mestách často riešený pomocou platených parkovísk, resp. parkovacích budov. V prípade akademickej pôdy je toto riešenie nielen finančne ale aj priestorov náročné. Preto sme na základe analýzy aktuálnych podmienok a dostupných technologických nástrojov navrhli prototyp elektronickej služby evidencie dočasného parkovania. Tento prototyp využíva metódy a prístupy počítačového videnia a umelej inteligencie. Pri výbere vhodnej predtrénovanej neurónovej siete sme uprednostnili kompromis medzi presnosťou klasifikácie a výpočtovou náročnosťou. Používateľské rozhranie umožňuje monitorovanie obsadenosti parkovacieho miesta v reálnom čase a notifikácie na príchod alebo odchod auta.

**Kľúčové slová** — Tensorflow, OpenCV, rozpoznávanie objektov

## I. ÚVOD

Správa parkovacích miest je v súčasnej dobe veľmi rozšírený problém najmä v mestách, kde sa rozvíjajú nové pracovné príležitosti a zároveň sa aktívne nebudujú nové sídliska takým tempom, aký je dopyt. Z toho vyplýva, že mesto ako napr. Košice budú mať vyššiu hustotu obyvateľstva, než tomu bolo v minulosti a tým sa budú zvyšovať aj počet áut, ktoré budú v meste parkovať, kde bude nevyhnutnosťou nejakým spôsobom tieto parkovacie miesta regulovať. Zvýšením tlakom po dopyte parkovacích miest v rámci mesta bude priamo úmerne sa zväčšovať aj tlak na využitie parkovacích miest v rámci akademickej pôdy TUKE, kde parkovacie miesta nie sú zatiaľ až tak prísne regulované, čo sa v blízkej perspektíve zmení.

Tieto problémy sú v mestách riešené pomocou platených parkovísk, resp. parkovacích budov. Ale samostatné inštitúcie ako napr. vyššie spomínané univerzity nemajú také možnosti a vyhradený personál ako samotné parkovacie domy, parkovacie miesta spravované mestom a pod. Mnohokrát kvôli parkovnému vodiči zneužívajú takéto inštitúcie (univerzity, školské internáty a pod.) na parkovanie svojich aut, aby sa vyhli plateniu. Tým na druhej strane zaberajú parkovacie miesta tým, čo by mali mať priestor na parkovanie ako napr. zamestnanci, hostia, študenti a iní, čo sú nejakým spôsobom zainteresovaní na týchto miestach. Inteligentné parkovacie systémy, ktoré sú k dispozícii v súčasnej dobe, si vyžadujú veľký investičný vstup, lebo firmy špecializujúce sa na takéto riešenia dodávajú hotový produkt s využitím ich IT infraštruktúry. Tento prístup je veľmi rentabilný pre riešenia parkovacích budov a iných platených parkovacích zariadení, ale ekonomicky nevýhodné pre univerzity alebo iné školské zariadenia.

Keďže TUKE ako aj iné univerzity, či iné vysoké školy nemajú kapacity na to, aby si platili špeciálnych zamestnancov pre správu parkovacích miest, tak sú nútení tieto úlohy delegovať aktuálnym zamestnancom, napríklad vrátnikom. Títo zamestnanci majú aj inú náplň ako kontrolu parkovacích miest a samozrejme, nie každý zo zamestnancov má právomoc na vznikajúcu problémovú situáciu nejakým spôsobom zasahovať. Na základe týchto problémov vzniká situácia, kde je potrebné analyzovať vnútro organizačné procesy, navrhnúť optimalizáciu týchto procesov a samozrejme implementovať hardvérovo-sofтверové riešenie.

## II. VÍZIA RIEŠENIA

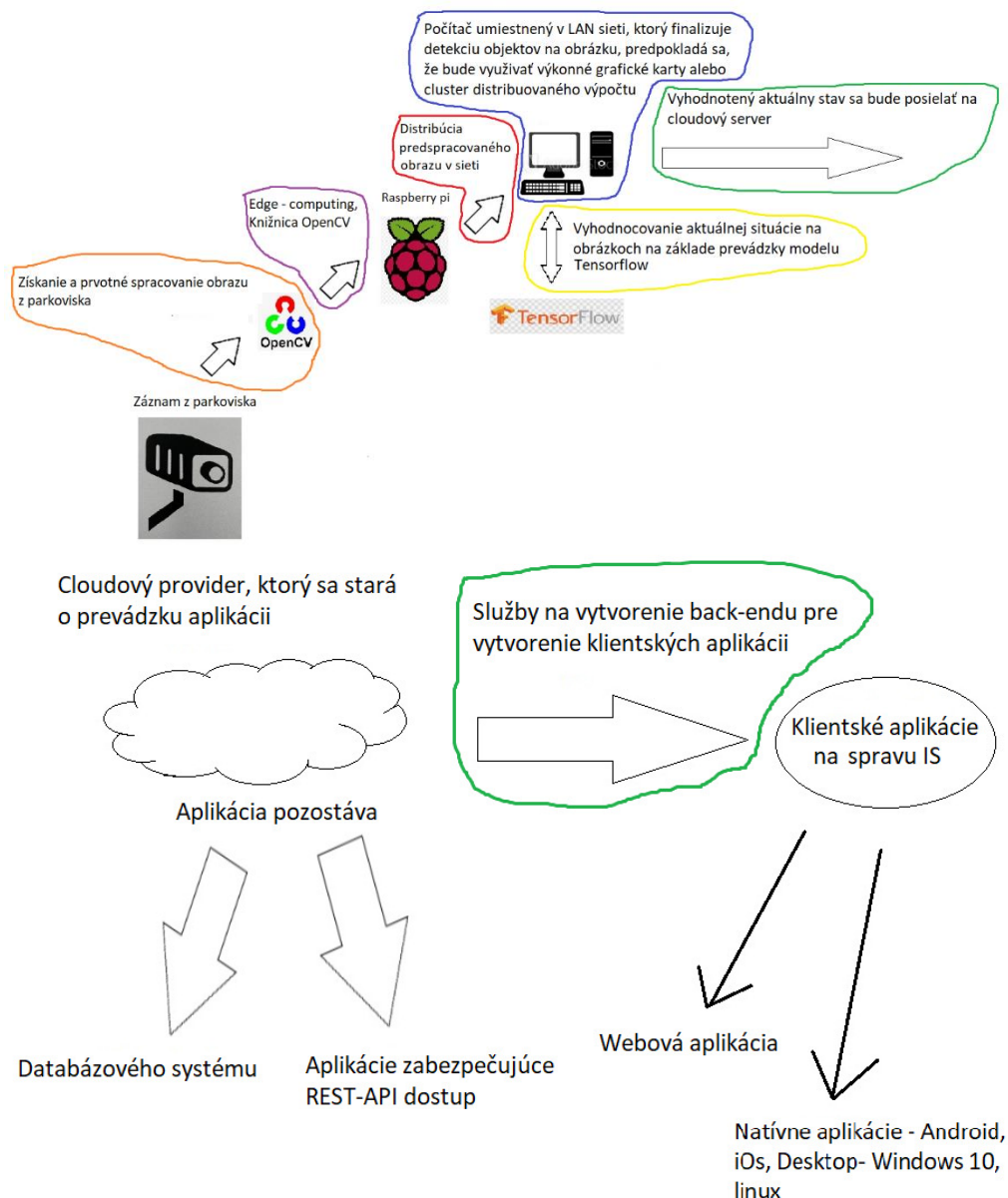
Riešením je prototyp jednoduchého informačného systému (IS), ktorý ušetrí náklady na potenciálneho zamestnanca navyše, keďže monitorovanie, zber a jednoduchá analýza dát bude plne automatizovaná. Kompetentná osoba bude môcť na základe automatických notifikácií

prijímať riešenia podľa organizáciou stanovených interných predpisov.

Host' zaparkuje auto na dočasnem parkovacom mieste, na základe čoho IS pošle notifikáciu zamestnancovi TUKE a zároveň začne odpočítavať čas na to, aby bolo parkovisko využité podľa predpisov TUKE. Akonáhle sa čas prekročí, tak zamestnanec TUKE osloví majiteľa vozidla, aby parkovisko opustil, prípadne ak nepozná majiteľa vozidla, tak bude postupovať podľa interných predpisov TUKE. Ak majiteľ včas opustí parkovacie miesto, tak sa pošle notifikácia zamestnancovi TUKE, že parkovacie miesto je voľné a pripravené pre iného host'a.

### III. ARCHITEKTÚRA

Architektúra IS je popísaná na obrázkoch nižšie. Rozdelili sme ju na back-endovú časť a na klientsku časť, kde využívame aj edge computing aj cloudové riešenia. Databáza MongoDB je na strane cloudu a ostatné časti sú v našom prípade nasadené na edge computingu. Raspberry Pi s kamerovým modulom za pomoci knižnice OpenCV [4] vysiela do iného počítača v lokálnej sieti, v našom prípade bol to notebook značky lenovo Z50-70 využívajúcu procesor značky Intel Core i5 s taktovacou frekvenciou 1,70 GHz po maximálnu taktovaciu frekvenciu 2,40 GHz.



Obr. 1 Architektúra navrhnutého IS

Raspberry Pi spracuje obraz (tyrkysová farba na obr.1) a posielala ho vo forme maticového výstupu do notebooku (červená farba), kde beží model neurónovej siete (modrá a žltá farba). Výsledky sa posielajú do cloudu a klientskych aplikácii (zelená farba).

V praxi sa rozpoznávanie objektov v obrazoch veľmi často používa v rôznych doménach, preto nie je efektívne pre každú aplikáciu zvlášť navrhovať novú neurónovú sieť. Typický postup

je použiť osvedčené architektúry neurónových sietí z jednej domény a transformovať ich do inej situácie a podmienok. Tomuto prístupu hovoríme „transfer learning“ (v prekl. „prenesené učenie“) podľa[1]. Funguje tak, že jednotlivé váhy sa vo všetkých vrstvách prenášajú, len posledná vrstva klasifikačná (dopredná) vrstva sa odstráni a vytvorí sa nová, v ktorej je počet neurónov identických ako počet tried, ktoré chceme klasifikovať. Potom sa váhy v poslednej vrstve dotrénujú na vzorke anotovaných obrázkov tých objektov, ktoré chceme vo svojej aplikácii za pomoci neurónových sietí hľadať. Voľne dostupných architektúr je momentálne veľa a výber závisí od toho, či chceme, aby bola aplikácia rýchla a sme ochotní pripustiť jej menšiu presnosť a návratnosť. Alebo máme záujem o spoľahlivý klasifikátor na úkor rýchlosti výpočtu klasifikácie. Architektúra na obr.2 je vyvážením presnosti modelu a rýchlosti klasifikácie.

Type / Stride	Filter Shape	Input Size	
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$	
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$	
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$	
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$	
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$	
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$	
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$	
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$	
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$	
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$	
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$	
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$	
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$	
5×	Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
	Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$	
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$	
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$	
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$	
Avg Pool / s1	Pool $7 \times 7$	$7 \times 7 \times 1024$	
FC / s1	$1024 \times 1000$	$1 \times 1 \times 1024$	
Softmax / s1	Classifier	$1 \times 1 \times 1000$	

### Whole Network Architecture for MobileNet

Obr. 2 Architektúra predtrénovanej neurónovej siete v softvérovej knižnici Tensorflow.podľa [3][2]

#### IV. TESTOVANIE

V návrhu IS sme sa zamerali na schopnosť vyhodnocovať situáciu na parkovisku v reálnom čase. Tento aspekt bol aj hlavnou náplňou realizovaného testovania, t.j. dosiahnuť správnu klasifikáciu sledovaného objektu (auta) z vysielaného videa a čo najaktuálnejšie a najplynulejšie klasifikované video na výstupe. Prvým krokom bolo dosiahnuť minimálnu hodnotu 30 FPS (frames per second, z angličtiny snímky za sekundu). Hodnota 30 FPS je minimum na to, aby ľudské oko vnímalo zmenu obrázkov ako plynulý jav a 60 FPS je odporúčaný stav. Z dôvodu výpočtovej náročnosti je 60 FPS veľmi vysoká hodnota, ktorá sa ťažko dosahuje, lebo je to náročné pre dátový tok aj výpočtové kapacity počítačov. Dosiahnuť 30 FPS je ale splniteľnou hodnotou.





Obr. 3 Vyhodnotenie FPS a presnosť klasifikácie auta označeného v zelenom obdĺžniku.

Na obr.3 je uvedená presnosť klasifikácie objektu pri ideálnych svetelných podmienkach a v ľavom hornom rohu je uvedený aktuálny prepočet FPS streamovaného videa z webkamery pripojenej k Raspberry Pi.

## V. ZÁVER

Cieľom bolo vytvoriť prototyp elektronickej služby evidencie dočasného parkovania na danom parkovisku. Na riešenie tohto problému sme sa rozhodli použiť metódy a prístupy počítačového videnia a zber analyzovaného videa za pomoci kamier. Algoritmy počítačového videnia a hlbokých konvolučných neurónových sietí nám umožnili vytvoriť plno-automatizovaný prototyp. V budúcnosti, ak sa tento prototyp osvedčí v rámci reálneho prostredia, pre ktoré bol navrhnutý, tak sa otvorí priestor pre jeho nasadenie a vylepšovanie aj na iných parkoviskách.

Z pohľadu využívania metód umelej inteligencie na riešenie rôznych životných situácií konštatujeme, že konvolučne neurónové siete vhodnej architektúry v súčasnej dobe dominujú v oblastiach zložitých problémoch počítačového videnia nad metódami a algoritmi, ktoré nepoužívajú neurónové siete. Týka sa to výkonnosti aj presnosti riešenia jednotlivých problémov. Nevýhodou je, že na druhej strane sú veľmi nepraktické na veľmi jednoduché úlohy, ktoré sa v praxi takmer vôbec nevyskytujú.

## ZDROJE

- [1] A. Rosebrock, "Deep Learning for Computer Vision with Python", vol. 53. 2017.
- [2] Simplilearn, *Top 10 Deep Learning Frameworks You Should Know in 2021*. Dostupné: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-frameworks>
- [3] S. Tsang, *Review: MobileNetV1-Depthwise Separable Convolution (Light Weight Model)*. Dostupné: <https://towardsdatascience.com/review-mobilenetv1-depthwise-separable-convolution-light-weight-model-a382df364b69>
- [4] R. Yadav, *Top 7 Image Processing Libraries In Python*. Dostupné: <https://analyticsindiamag.com/top-8-image-processing-libraries-in-python/>

# Dielektrická spektroskopia transformátorového GTL oleja

<sup>1</sup>Peter Havran, <sup>2</sup>Roman Cimbala, <sup>3</sup>Maksym Oliinyk

<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

<sup>1</sup>peter.havran@tuke.sk, <sup>2</sup>roman.cimbala@tuke.sk, <sup>3</sup>maksym.oliinyk@tuke.sk

**Abstrakt** — Táto publikácia rieši správanie sa progresívneho elektroizolačného oleja Shell DIALA S4 ZX-1, vyrobeného na báze technológie GTL, v urýchlennom degradačnom procese. V úvode je popísaná dôležitosť izolačného systému transformátorov a metóda dielektrickej spektroskopie na vyšetrovanie dejov, prebiehajúcich v skúmanom materiáli. Experiment obsahuje popis metodiky merania, meracieho pracoviska a analýzu experimentálnych dát vo frekvenčnej oblasti dielektrickej spektroskopie.

**Kľúčové slová** — dielektrická spektroskopia, komplexná permitivita, olej, technológia GTL

## I. ÚVOD

Dôležitosť izolačného systému spočíva v garancii bezproblémového prevádzkovania elektroenergetického zariadenia na každej napäťovej úrovni. Preto je potrebné sledovať stav zariadenia a jeho izolačného systému, podliehajúceho nenávratným fyzikálno-chemickým zmenám v podobe starnutia. V dôsledku starnutia výkonových transformátorov je dnešný elektroenergetický priemysel zaťažovaný veľkou výzvou, pretože havárie transformátora a jeho následná oprava je dosť nákladná. Transformátory predstavujú jeden z najdôležitejších zariadení v elektrizačnej sústave, preto potreba spoľahlivých metód monitorovania a vyšetrovania stavu izolačného systému poháňa uznávaných svetových odborníkov k vývoji a návrhu nových technológií, ktoré zlepšujú spoľahlivosť prevádzky a optimalizujú chod všetkých komponentov siete [1]. Jednou z progresívnych návrhov v novodobom priemysle je transformátorový inhibovaný elektroizolačný olej Shell DIALA S4 ZX-1 od spoločnosti Shell, ktorý je dimenzovaný na zvládnutie náročných výziev najnovších výkonových transformátorov. Tento olej disponuje predĺženou životnosťou s dôrazom na nulový obsah síry vďaka výrobnéj technológii GTL, ktorá predstavuje unikátny proces výroby syntetického základového oleja, premieňajúceho zemný plyn na kvapalné uhl'ovodíky, výsledkom čoho je priezračne čistý základový olej [2]. O jeho unikátnosti niet z hľadiska chemických parametrov pochyb, ako je známe z jeho preferencií, preto je potrebné získať čo najviac informácií o fyzikálno-chemických vlastnostiach v podobe komplexného vyšetrovania dejov, prebiehajúcich v oleji pri starnutí s cieľom jeho bezproblémovej aplikácie v praxi. V tejto publikácii bude skúmaná zmena správania sa GTL oleja Shell DIALA S4 ZX-1 (ďalej len „GTL-SD“) pri urýchlennom tepelnom starnutí prostredníctvom metódy dielektrickej spektroskopie vo frekvenčnej oblasti (FDS).

Technológia dielektrickej spektroskopie má potenciál poskytnúť správne hodnotenie stavu dielektrických materiálov, ktoré sú široko používané na celom svete a nahrádza preto iné diagnostické metódy. Metódy dielektrickej spektroskopie zahŕňajú metódu zotaveného napätia (RVM) a metódu polarizačného a depolarizačného prúdu (PDC), ktoré sú založené na reakcii voľných a viazaných nábojov pod vplyvom jednosmerného elektrického poľa v časovej oblasti. Metóda dielektrickej spektroskopie vo frekvenčnej oblasti je založená na reakcii vo frekvenčnom pásme, ktorá využíva polarizačnú charakteristiku dielektrika pri striedavom elektrickom poli na výpočet komplexnej permitivity  $\mathcal{E}^*$  a ostatných z nej odvodených parametrov, ako je kapacita  $C$  a stratový činiteľ  $tg \delta$  ako funkcie frekvencie. V porovnaní s metódou RVM a PDC má metóda FDS dobrú výkonnosť filtrovania šumu, najmä pri vysokej

frekvencii, ktorá predstavuje prioritu pre skúmanie materiálov s nízkymi dielektrickými stratami [3].

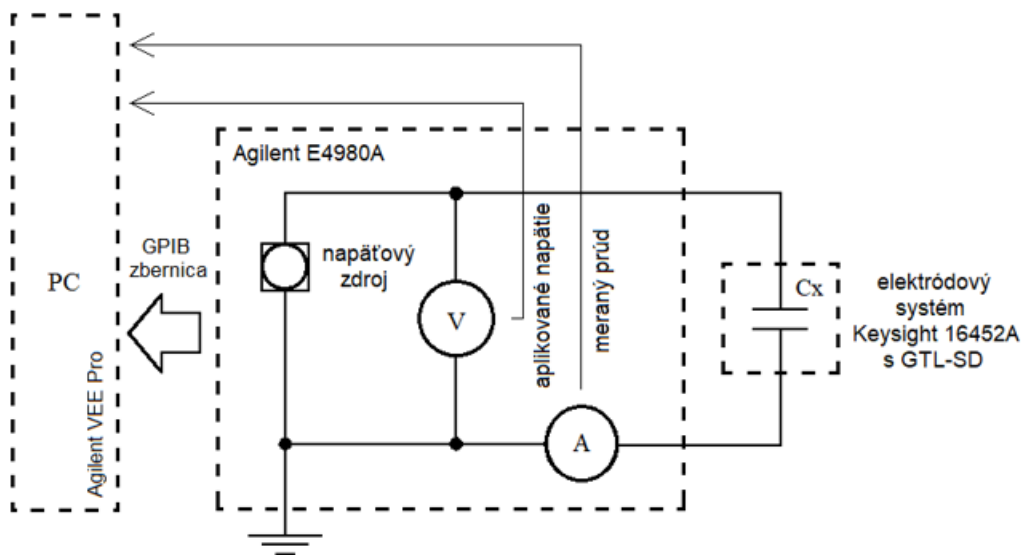
## II. EXPERIMENT

Princípom tohto experimentu bolo sledovanie zmien a porovnanie dielektrických vlastností elektroizolačného oleja GTL-SD, podliehajúceho urýchlenému tepelnému starnutiu, ktoré pozostávalo zo siedmich meraní pri časovej periodicite 250 h, 430 h, 600 h, 814 h, 1080 h, 1580 h a 1830 h. Všetky merania boli realizované vo frekvenčnej oblasti dielektrickej spektroskopie, pričom boli získané hodnoty kapacity a stratového činiteľa, ktoré boli následne prepočítané na hodnoty komplexnej permitivity, teda na jej reálnu a imaginárnu zložku s cieľom získania komplexného správania relaxačných mechanizmov v skúmanom elektroizolačnom oleji.

Merania boli realizované na novej nepoškodenej vzorke elektroizolačného oleja GTL-SD. Príprava pred tepelným starnutím zahŕňala odobratie oleja do sklenenej fľaše, do ktorej bol vložený zvinutý kus medi o dĺžke 20 cm a priereze  $2,5 \text{ mm}^2$ . Význam medi v skúmanej vzorke spočíval v jej chemických vlastnostiach, keďže meď pôsobí ako katalyzátor pri chemických reakciách v elektroizolačnom oleji počas starnutia.

Meranie bolo uskutočňované v profesionálnom laboratóriu pri izbovej teplote a atmosférickom tlaku vzduchu. Urýchlené tepelné starnutie bolo zabezpečované pomocou univerzálnej rúry UF55<sup>plus</sup>, ktorá sa využíva na vykurovanie, sušenie a starnutie vo výskume, vede a priemysle. Fľaša s GTL-SD bola uložená do rúry, ktorá bola nastavená na teplotu  $90 \text{ }^\circ\text{C}$ . Meranie pred začiatkom tepelného starnutia (0 h) nebolo realizované kvôli odhadovanej vysokej vlhkosti oleja, pretože nebolo známe, kde bol GTL-SD skladovaný.

Pred každým meraním, po vytiahnutí z rúry, vzorka GTL-SD chladla jeden deň z dôvodu uskutočnenia merania. Medzitým bol poskladaný elektródový systém Keysight 16452A, do ktorého bola po vychladnutí aplikovaná vzorka. Zapojenie meracieho pracoviska na Obr. 1 obsahovalo elektródový systém s GTL-SD, prepájacie vodiče, merací prístroj a laptop. Meranie bolo realizované meracím prístrojom Agilent E4980A, ktorý bol prepojený cez GPIB zbernicu na laptop s riadiacim softvérom Agilent VEE Pro, zaznamenávajúcím namerané dáta stratového činiteľa a kapacity pri nastavenom striedavom napätí 2 V.



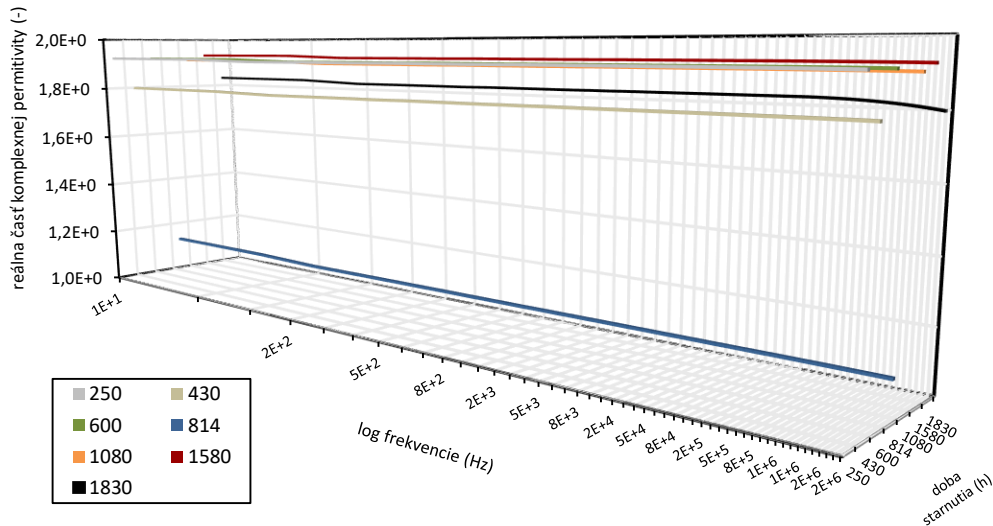
Obr. 1 Zapojenie meracieho pracoviska

## III. ANALÝZA EXPERIMENTÁLNYCH DÁT

Táto kapitola zahŕňa analýzu experimentálnych dát reálnej a imaginárnej časti komplexnej permitivity pri odstupňovanej dobe tepelného starnutia (250 h, 430 h, 600 h, 814 h, 1080 h, 1580 h a 1830 h) vo frekvenčnom spektre LCR metra od 10 Hz do 2 MHz pri aplikovanom napätí 2 V.

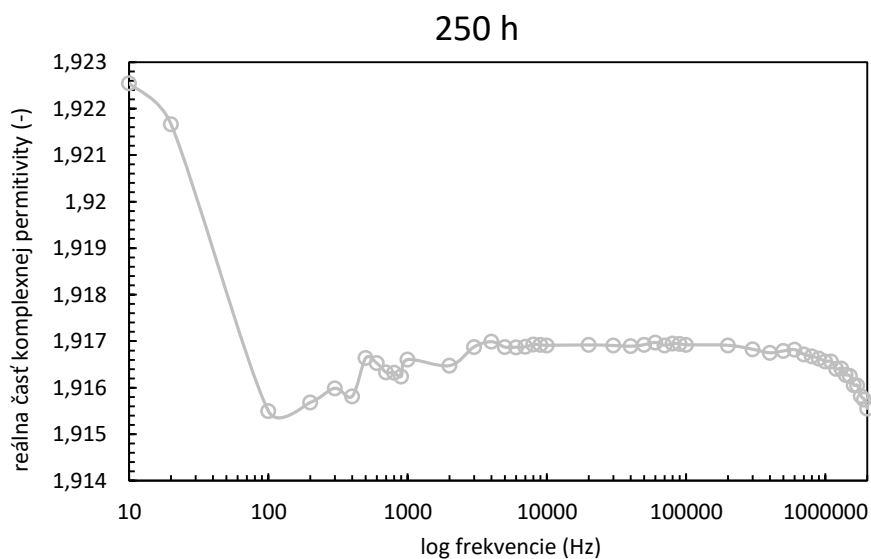
Na Obr. 2 je znázornené porovnanie reálnej časti komplexnej permitivity (ďalej len „reálnej permitivity“) pri odstupňovanej dobe tepelného starnutia. Z porovnaní charakteristík je vidieť dynamicky sa meniace správanie kapacitných pomerov pri zvyšujúcej sa dobe tepelného starnutia. Medzi dobou tepelného starnutia 250 h a 430 h došlo k poklesu reálnej permitivity

o 6,5% a následne k zvýšeniu o 6,3% pri starnutí 600 h. Následne došlo k výraznému poklesu pri 814 h o 41,3% a spôsobilo to zníženie prehľadnosti porovnania grafických závislostí pri rôznej dobe tepelného starnutia. Pri dobe 1080 h stúpila reálna permitivita približne na pôvodné hodnoty o 41%. Tepelné strnutie 1580 h spôsobilo mierne zvýšenie hodnôt o 1% a pri 1830 h klesla reálna permitivita o 5,1% oproti dobe 1580 h. Z týchto porovnaní nie je možné bližšie charakterizovať tento dynamicky sa meniaci trend. Ak si však porovnáme tepelno-degradačnú závislosť bez charakteristík pri 430 h, 814 h a 1580 h, teda porovnaním pri 250 h, 600 h, 1080 h a 1830 h, možno dôjsť k záveru, že zvýšením doby tepelnej degradácie dochádza k zníženiu reálnej permitivity oleja GTL-SD, resp. k zníženiu pomeru sily pôsobiacej na elektrický náboj v danom dielektrickom materiáli podľa Coulombovho zákona.



Obr. 2 Závislosť reálnej časti komplexnej permitivity na frekvencii pri rôznej dobe tepelného starnutia

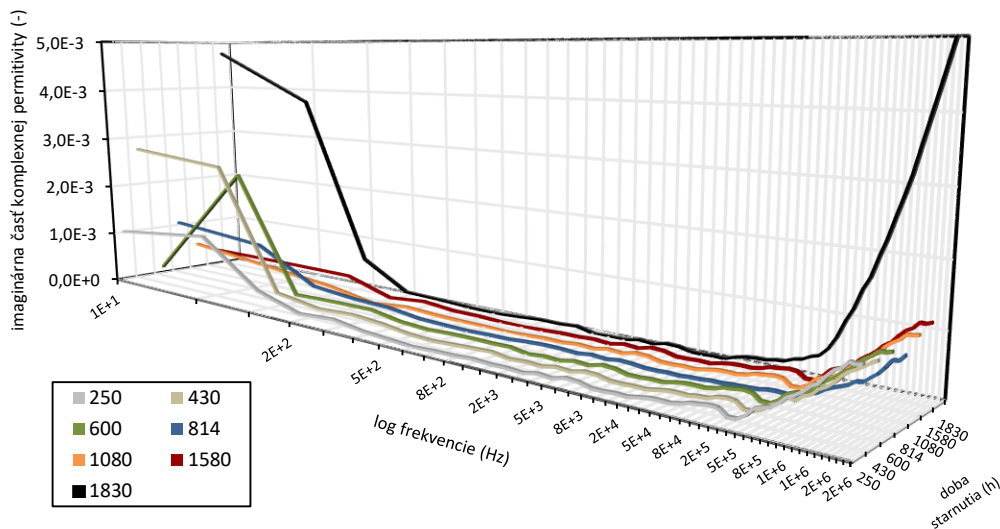
Vo frekvenčnom spektre od 10 Hz do 2 MHz boli zaznamenané dva relaxačné procesy, ktoré nie sú z Obr. 2 zachytené v dôsledku vyššieho rozsahu hodnôt reálnej permitivity na osi y v degradačnom procese. Vynesením každej závislosti pri určitej tepelnej degradácii zvlášť, bádať uvedené polarizačné deje vo frekvenčnom pásme od 10 Hz do 100 Hz a od 600 kHz do 2 MHz, ako je to znázornené na Obr. 3 pri tepelnej degradácii 250 h. Uvedené polarizačné procesy boli zaznamenané v rovnakých frekvenčných pásmach pri každej dobe tepelnej degradácie, avšak s inou hodnotou reálnej permitivity, odrážajúcou veľkosť dipólových momentov molekúl v skúmanom materiáli.



Obr. 3 Reálna časť komplexnej permitivity v závislosti od frekvencie pri tepelnom starnutí 250 h

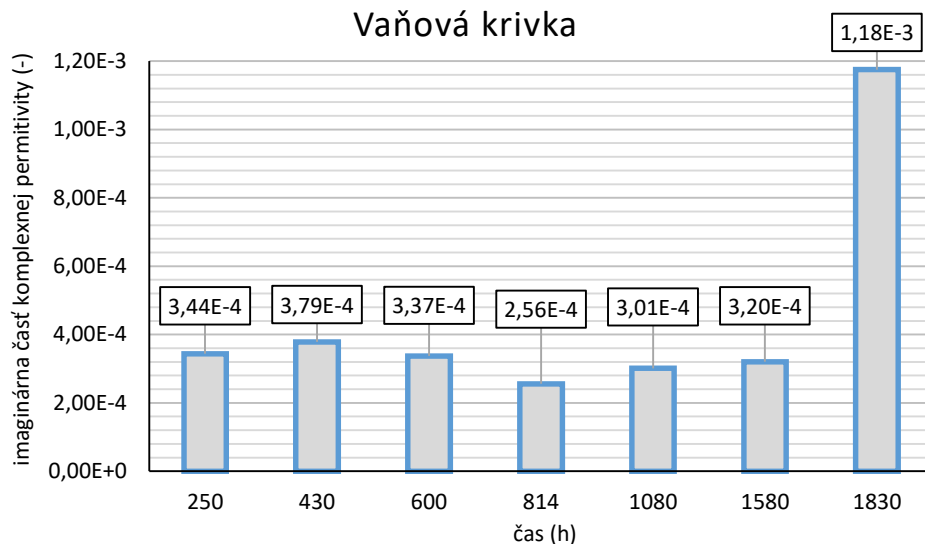
Na Obr. 4 je znázornené porovnanie imaginárnej časti komplexnej permitivity (ďalej len „imaginárnej permitivity“) pri odstupňovanej dobe tepelného starnutia. Imaginárna permitivita predstavuje v komplexnej rovine materiálu jeho dielektrické straty. Vo frekvenčnom pásme

meracieho prístroja od 10 Hz do 2 MHz boli podobne ako pri reálnej permitivite zaznamenané dva nedostatočne rozvinuté polarizačné mechanizmy, ktoré spôsobili zvýšenie imaginárnej permitivity v nízkofrekvenčnej a vo vysokofrekvenčnej oblasti. Nízkofrekvenčný proces môže byť spojený s polarizáciou iónovej nečistoty, rozptýlenej v oleji GTL-SD. V oblasti stredných frekvencií sa dielektrické straty pohybovali na približne rovnakej úrovni pri rôznom stupni tepelnej degradácie. Vo vysokofrekvenčnej oblasti došlo k zaznamenaniam relaxačného procesu, pripisovaného orientačnej (dipólovej) polarizácii prostredníctvom zvýšenia dielektrických strát, kedy dochádza k ohrevu materiálu, spôsobeného oneskoreným natáčaním dipólov za zmenami frekvencie vonkajšieho elektrického poľa.



Obr. 4 Závislosť imaginárnej časti komplexnej permitivity na frekvencii pri rôznej dobe tepelného starnutia

Z grafickej závislosti na Obr. 4 nedochádzalo k postupnému zvyšovaniu dielektrických strát s narastajúcou dobou tepelného starnutia. Mierny nárast s postupným tepelným namáhaním je badateľný v oblasti vyšších frekvencií pri rozvíjaní orientačnej polarizácie, kedy dochádza vplyvom tepelnej degradácie k oxidačným procesom, ktorých výsledkom sú kyslé a iné polárne zlúčeniny, zapríčiňujúce zvýšenie dielektrických strát. Zo všeobecného hľadiska je možné povedať, že správanie dielektrických strát GTL-SD vykazuje pri urýchlennom tepelnom starnutí 1830 h pri teplote 90 °C takzvanú vaňovú krivku, ktorá sa výrazne prejavila medzi tepelným starnutím 1580 h a 1830 h, kedy došlo k nárastu dielektrických strát o 268,8%. Uvedená vaňová krivka je znázornená na Obr. 5, pričom jej dáta boli získané spriemerovaním hodnôt imaginárnej permitivity pri každom stupni tepelného starnutia v celom frekvenčnom spektre. Vaňová krivka je charakteristická tvarom priebehu do tvaru písmena U, kde so začiatkom procesu starnutia prebieha takzvané dotvrdzovanie materiálu so znížením dielektrických strát, resp. vylepšovaním jeho dielektrických vlastností. Pri dlhšej dobe starnutia dochádza k tvorbe oxidačných procesov, spojených s uvoľňovaním polárnych látok, ktoré sú zodpovedné za zvýšenie dielektrických strát, smerujúcich k postupnému poškodeniu materiálu až k jeho havárii.



Obr. 5 Vaňová krivka elektroizolačného oleja GTL-SD

#### IV. ZÁVER

Cieľom tohto experimentu bolo sledovanie zmien a porovnanie dielektrických vlastností elektroizolačného oleja GTL-SD, podliehajúceho urýchlenému tepelnému starnutiu. Porovnaním pri 250 h, 600 h, 1080 h a 1830 h, možno dôjsť k záveru, že zvýšením doby tepelnej degradácie dochádza k zníženiu reálnej permitivity oleja GTL-SD. Vo frekvenčnom pásme meracieho prístroja od 10 Hz do 2 MHz boli podobne ako pri reálnej permitivite zaznamenané dva nedostatočne rozvinuté polarizačné mechanizmy, ktoré spôsobili zvýšenie imaginárnej permitivity v nízkofrekvenčnej a vo vysokofrekvenčnej oblasti. Z grafickej závislosti na Obr. 4 nedochádzalo k postupnému zvyšovaniu dielektrických strát s narastajúcou dobou tepelného starnutia. Mierny nárast s postupným tepelným namáhaním je badateľný v oblasti vyšších frekvencií. Zo všeobecného hľadiska je možné povedať, že správanie dielektrických strát GTL-SD vykazuje pri urýchlenom tepelnom starnutí 1830 h pri teplote 90 °C takzvanú vaňovú krivku, ktorá sa výrazne prejavila medzi tepelným starnutím 1580 h a 1830 h, kedy došlo k nárastu dielektrických strát o 268,8%.

#### POĎAKOVANIE

Práca bola podporená MŠVVaŠ v rámci projektu VEGA 2/0011/20 a 1/0154/21 a Slovenskou agentúrou pre výskum a vývoj na základe zmlúv č. APVV-15-0438, APVV-17-0372 a APVV-18-0160.

#### LITERATÚRA

- [1] B. P. Pokharel, N. R. Karki, „Electrical Engineering Materials“, Oxford, U. K.: Alpha Science International Ltd. 2007, 214 pp. ISBN-13:978-1-84265-297-8.
- [2] Shell DIALA Transformers Oils. Online: <<https://www.shell.com/promos/lubes/btb-products/diala-product-range>>.
- [3] J. Liu, Y. Xu, X. Wei, X. Li, „Research of Dielectric Spectroscopy on Insulation Ageing Assessment of XLPE Cables“, IEEE, 2013 Annual Report Conference on Electrical Insulation and Dielectric Phenomena. 2014, pp. 140-143, ISBN 978-1-4799-2597-1.

# Využitie open-source platformy na elektroenergetické účely

*Patrik HRIVNIAK, Dušan MEDVEĎ*

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

patrik.hrivniak@student.tuke.sk, dusan.medved@tuke.sk

**Abstrakt** — Tento článok je venovaný návrhu modelu domáceho elektromera, ktorého funkcionality zahŕňa meranie, zaznamenávanie a analyzovanie základných elektrických veličín v domovej elektroinštalácii. Pomocou tohto zariadenia boli získané a následne vyhodnotené namerané údaje vybraných elektrických veličín.

**KLúčové slová** — Arduino, elektroinštalácia, smart-sieť, meranie, smart elektromer

## I. ÚVOD

Touto témou sme sa rozhodli zaoberať pretože cieľom tohto príspevku bolo navrhnúť jednoduchý model domáceho elektromera ako alternatívy voči konvenčným, bežne dostupným elektromerom na trhu. Pomocou nášho modelu je možné monitorovať, zaznamenávať a následne analyzovať základné parametre a elektrické veličiny v domovej elektroinštalácii za určité časové obdobie.

V tomto článku si najskôr uvedieme súčiastky potrebné na skonštruovanie nášho zariadenia, algoritmus merania údajov a algoritmus spracovania údajov. Ďalej si ozrejmime, aké je základné elektrické zapojenie, aké veličiny budeme sledovať a na záver verifikujeme namerané výsledky.

## II. ELEKTROMER PRE DOMÁCE VYUŽITIE

### A. Fyzické súčiastky a zostavenie modelu

Pre náš projekt sme si zvolili open-source platformu Arduino. Na oficiálnych stránkach spoločnosti Arduino [1] a [2] je v krátkosti opísaná táto voľne dostupná elektronická platforma, využívajúca pomerne ľahko ovládateľný hardvér a softvér. Táto platforma je vyhotovená vo forme prototypovacej dosky s mikroprocesorom, pomocou ktorého môžeme riadiť, kontrolovať a komunikovať so širokou škálou senzorov, displejov, motorčekov a ostatných periférií. Vieme spracovávať vstupy a výstupy, rôzne druhy údajov zo senzorov a podobne. Táto komunikácia prebieha vo vývojovom prostredí Arduina pomocou vlastného programovacieho jazyka. Arduino vie teda prepojiť rôzne menšie aj väčšie súčiastky do funkčného celku a tvorí tak základ rozmanitých projektov.

Ďalej na týchto webových stránkach sú uvedené rôzne varianty tejto platformy, ako aj široký výber produktov od samotných dosiek cez rôzne moduly, shieldy až po zostavy viacerých produktov. Rôzne produkty môžu mať podobné, ale aj veľmi rozličné využitie. Dôležitým faktorom v rozhodovacom procese pri kúpe produktu bol fakt, že schémy potrebné na skonštruovanie samotných dosiek sú podobne, ako všetko ostatné týkajúce sa Arduina, voľne dostupné, teda prakticky každý má, za určitých podmienok, možnosť si takúto dosku vyrobiť. Je teda na každom zvážiť plusy a mínusy týchto možností a vybrať si produkt, ktorý bude jednotlivým ľuďom a jednotlivým projektom vyhovovať najviac. V našom prípade sme zvolili poslednú spomínanú variantu, verziu Arduino UNO, ktorá je vhodná dostupnosťou zdrojov a návodov, ako aj kapacitou a funkcionalitou.

Ako meracie zariadenie sme použili modul PZEM 004-T v3.0 (obr. 1).



Obr. 1 Modul PZEM 004-T v3.0

Toto zariadenie je možné využiť na meranie rôznych elektrických veličín striedavého prúdu a napätia v domácnosti. Medzi ďalšie veličiny, ktoré toto zariadenie meria patrí napätie, prúd, výkon, účinník, frekvencia, či spotreba elektrickej energie. Všetky tieto veličiny je schopné merať v určitom rozsahu pri určitej odhadovanej presnosti. Ďalej sme sa v oficiálnom obchode spoločnosti „Peacefair“ v [3] dozvedeli o rôznych typoch meracích zariadení, z ktorých sme si vybrali pre naše účely vhodné meracie zariadenie. Pri výbere sme uvažovali dostupnosť na trhu a využiteľnosť. K zariadeniu sme potrebovali použiť externý transformátor prúdu, kde sme si ďalej mohli vybrať rozpojovací alebo transformátor vyhotovený v jednom celku. Tieto externé prístroje zvládnu záťaž do 100 A, teda boli vhodné pre náš projekt. Rovnako bol pre nás vhodnejší transformátor rozpojovací, kvôli jednoduchšej manipulácii pri zapájaní a prípadnom odpájaní, či údržbe zariadenia. Zariadenie je určené na prácu pri  $-20$  až  $+60$  °C a využíva RS485 komunikačné rozhranie pri rýchlosti 9600 baudov. Spojenie s Arduinoom je realizované cez sériovú TTL komunikáciu.

Ďalším prvkom obvodu bol LCD displej, teda vizualizačné zariadenie pre Arduino, nakoľko bolo potrebné zobrazovať rôzne údaje či komunikovať s užívateľom. Poznáme rôzne typy a veľkosti displejov, pričom pre náš projekt sme si vybrali 4 riadkový a 20 znakov široký displej, pretože chceme zobrazovať viacero údajov naraz a cenovo nie je medzi menším a často využívaným dvoj-riadkovým modelom veľký rozdiel. Podobne prízvukuje využitie LCD displeja na komunikáciu so zariadením aj Selecký v [4], kde ďalej opisuje skladbu displeja, ktorý je zložený buď z malých segmentov alebo tekutých krištáľov. Analogicky fungujú všetky LCD displeje, bez ohľadu na ich veľkosť, keďže na najmenšej úrovni sa skladajú z rovnakých zložiek.

Dôležitou funkcionalitou nášho modelu je záznam údajov. K tomu je potrebné presné zaznamenávanie času, na čo sme použili RTC teda „Real Time Clock“ modul. Autor publikácie [4] uvádza podrobnosti aj tohto modulu. Zariadenie predstavuje jednoduchý spôsob presného sledovania času od prvotného zadania. Je schopný zaznamenávať sekundy, minúty, hodiny a podobne aj dátum, teda dni, mesiace a roky. Je napájaný malou 3 V batériou a má nízku spotrebu energie. Tieto moduly sú teda vhodné do malých, prenosných zariadení, v ktorých je potrebné presné meranie času.

Ako úložisko pre merané elektrické veličiny sme sa rozhodli použiť SD kartu. Aby bolo Arduino schopné pracovať so spomínanou záznamovou SD kartou, bol potrebný na komunikáciu s ňou adaptér, vďaka ktorému je možné na ňu zapisovať želané údaje. Autori článku [5] použili rovnaký modul pre prácu s SD kartou v podobnom projekte. Uviedli o ňom: „Tento micro SD kartový modul je používaný na prenos dát do a zo štandardnej micro SD karty. Pinout je priamo kompatibilný s Arduinoom a tak isto môže byť použitý s inými mikrokontrolérmi“. Výber tohto modulu prebiehal podľa podobných kritérií ako ostatné súčiastky, pozerali sme sa na cenu, ľahkú použiteľnosť a kompatibilitu. Na prenos dát do počítača bol použitý jednoduchý adaptér pre USB prístup k počítaču a na presun dát pre ďalšie spracovanie v programe Matlab.

Medzi ďalšie súčiastky je možné zaradiť spojovacie vodiče a spojovaciu dosku, pomocou ktorých sú všetky ostatné väčšie súčasti spojené. Ďalej sme k týmto súčiastkam priradili adaptér na redukciu zásuvkového napätia na 9 V potrebných pre napájanie Arduina a z neho následne ďalších periférií. Tento variant sme uprednostnili pred použitím 9 V batérie (batéria by sa skoro vybila, keďže zariadenie je potrebné mať na mieste zapojené dlhú dobu, počas ktorej by nám batéria nevydržala). Použili sme preto jednoduchý, nastavovateľný adaptér na privedenie konštantného napájacieho napätia 9 V.

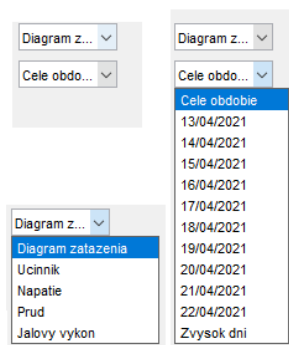
### B. Vytvorenie zdrojového kódu pre Arduino a Matlab

Meranie údajov prebieha vo všeobecnosti nepretržite od pripojenia a inicializácie SD karty. Prerušené je buď odpojením zariadenia od prívodu elektrickej energie alebo vybratím SD karty, kedy sa zariadenia opakovane každých 5 sekúnd rešartuje a kontroluje prítomnosť karty. V prípade jej nasledujúceho správneho vloženia sa pri najbližšom rešartovaní karta a súbor na



nej inicializujú a zahájí sa meranie s následným zapisovaním údajov. Jedna iterácia „loop“ slučky teda vyzerá približne nasledovne: aktualizácia času, kontrola prítomnosti SD karty, získanie aktuálnych hodnôt nameraných modulom PZEM, tieto hodnoty sa následne zapíšu na SD kartu do textového súboru spoločne s časom a dátumom a tiež sa zobrazia na LCD displej. Toto sa deje každú sekundu kvôli obmedzeniu vzorkovania samotného meracieho prístroja.

Pri tvorbe algoritmu spracovania našich nameraných údajov sme vychádzali z práce autorov Zaplatílek a Donar v [6]. Algoritmus sa opiera o niekoľko základných pilierov, ktoré sme chceli spracovaním z údajov zistiť. V tomto článku si uvedieme hlavný z nich – grafické znázornenie všetkých meraných veličín v určených časových horizontoch. Konkrétne sa jedná o zobrazenie celého sledovaného obdobia naraz, ktoré môžeme vidieť na obrázku 2. Ďalej sa jedná o zobrazenia jednotlivých dní, ako príklad si môžeme pozrieť obrázok 7. V programe je prednastavený počet dní, aký sa dá takto samostatne zobrazit'. V princípe je prvý deň vždy neúplný, ak teda spočítame len úplné dni (podľa našich meraní takto vieme zobrazit' 9 dní, počas ktorých sme realizovali merania pre daný článok). Zvyšok nameraných údajov je zobrazených v jednom celku kvôli praktickosti užívateľského rozhrania programu. Z týchto údajov sme usúdili, že ideálny čas merania, teda doba, počas ktorej je zariadenie zapojené a v jednom celku meria údaje, je približne 10 až 14 dní. Táto doba môže byť obmedzená kapacitou použitej pamäťovej karty, na ktorú sa dáta zaznamenávajú; je preto nutné použiť SD kartu s dostatočne veľkou kapacitou. Spomínané užívateľské rozhranie obsahuje dve možnosti výberu, v prvej si môžeme vybrať, akú meranú veličinu chceme zobrazit' a v druhej, za aké časové obdobie. Na nasledujúcom obrázku si môžeme všimnúť konkrétny príklad vyhotovenia tohto rozhrania.



Obr. 2 Užívateľské rozhranie výstupu z programu

### III. ELEKTROINŠTALÁCIA

Skôr než sa pozrieme konkrétne na domovú elektroinštaláciu, vysvetlíme si v skratke, ako sa dostaneme od verejného elektrického vedenia k rozvodu elektrickej energie v spomínanej elektroinštalácii. Toto prepojenie sa realizuje elektrickou prípojkou, ktorú autori v knihe [7] stručne opisujú takto: „Elektrická prípojka je elektrické vedenie, ktoré odbočuje od verejného rozvodu elektrickej energie smerom k odberateľovi a je určené na pripojenie odberných elektrických zariadení.“ Existujú rôzne druhy prípojok na základe vyhotovenia, či prevádzkového napätia, všetky sú však všeobecne tvorené odbočkou od verejného vedenia k prípojovej skrini. Z nej sa napája vedenie prechádzajúce vzduchom, zemou alebo kombinovane do elektromerového rozvádzača a z toho následne do hlavného domového rozvádzača. Podobný princíp je aj pri bytových rozvodniciach. Nás teda zaujíma hlavný domový, respektíve bytový rozvádzač, teda prívod od elektromerového rozvádzača. Tento prívod môže byť realizovaný z pohľadu káblového usporiadania dvoma spôsobmi, respektíve sústavami a tie sú TN-S a TN-C. Rozdielom medzi nimi bude počet vodičov, v prípade TN-C štyri a v prípade TN-S, ich bude 5 [7].

Pre nás sú podstatné tri fázové vodiče čiernej alebo hnedej farby. Potrebujeme ich na meranie prúdu a teda aj výkonu a spotreby elektrickej energie v domácnosti. Naše meracie zariadenie používa na meranie prúdu externý transformátor prúdu, ktorý na základe princípu elektromagnetickej indukcie v cievke meria veľkosť prúdu, ktorý prechádza vodičmi, ktoré prechádzajú cez cievku. Tento transformátor je jednoducho inštalovateľný na spomínané fázové vodiče, pretože sa dá otvoriť a zavrieť. Ak je v rozvádzači dostatok priestoru, je možné ho zapojiť bez nutnosti akéhokoľvek odpájania a ťažkej manipulácie s vodičmi. Pripojenie prvkov pre meranie napätia a ďalších veličín, je realizované ako ďalší paralelný zásuvkový obvod.

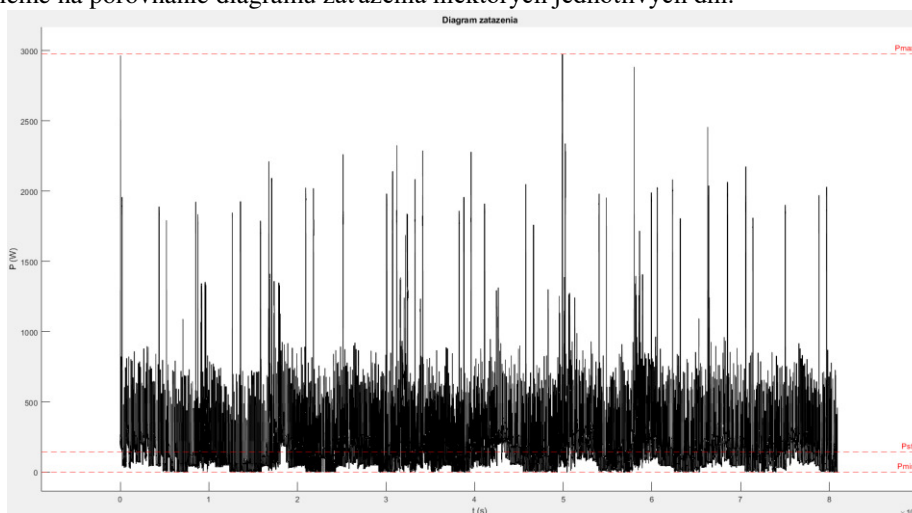
Opis týchto veličín sme čerпали prevažne od autorov publikácie [8], ktorí definujú elektrické napätie ako rozdiel dvoch elektrických potenciálov a podobne, elektrický prúd, ako tok elektrického náboja alebo tok elektrónov vodičom za určitý čas. Napäťová sústava, v ktorej je

dané zariadenie zapojené, je tvorená trojfázovým striedavým napätím s menovitou fázovou hodnotou 230 V a menovitou hodnotou združeného napätia 400 V s frekvenciou 50 Hz. S napätím úzko súvisia úbytky napätia definované ako rozdiel fázových napätí medzi dvoma uzlami, ktorý je tvorený tokom elektrického prúdu vetvou medzi nimi. Naše zariadenie ďalej meria činný a jalový výkon a spolu s nimi aj účinník. Elektrický výkon možno definovať ako prácu elektrického prúdu za jednotku času. V striedavých sústavách je reprezentovaný ako zdanlivý výkon  $S$ . Podstatnými sú pre nás jeho zložky: činný výkon a jalový výkon. Činný výkon predstavuje časť zdanlivého výkonu, ktorý sa premieňa na užitočnú prácu, teda zahrievanie tepelných spotrebičov, pohon elektromotorov a podobne. Jalový výkon reprezentuje energiu spotrebovanú na vytvorenie elektrického alebo magnetického poľa, teda energiu, ktorá sa nevyužila na užitočné napájanie elektrických spotrebičov. Podstatným ukazovateľom z hľadiska tejto premeny je účinník, ktorý vyjadruje mieru zdanlivého výkonu premeneného na činný výkon.

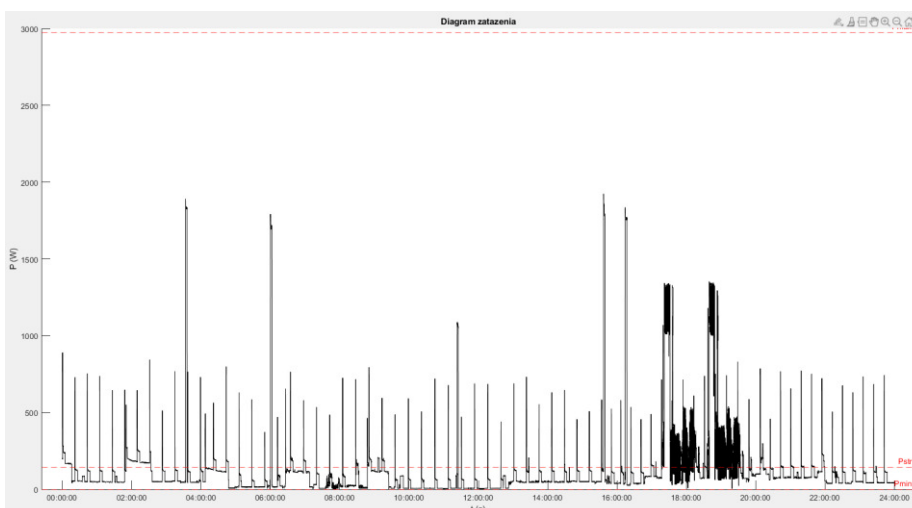
Na ilustráciu časovej postupnosti spotreby odoberaného činného výkonu sa často používa takzvaný denný diagram zaťaženia, ktorý: „vyjadruje časovú postupnosť zaťaženia v danej oblasti elektrizačnej sústavy alebo jej časti v priebehu daného dňa“ [8]. Používa sa hlavne pri analýze zaťaženia sústavy a odhade jej budúceho stavu. Jeho definíciu však môžeme použiť aj my na vyjadrenie nami nameraných hodnôt spotreby elektrickej energie v domácnosti.

#### IV. VÝSLEDKY MERANIA

Na obrázku 3 môžeme vidieť diagram zaťaženia vyobrazujúci činný výkon za celé sledované časové obdobie, t.j. približne 11 dní. Z tohto grafu si môžeme všimnúť hlavne maximálny dosiahnutý výkon 3000 W. K tejto krajnej hodnote sa údaje priblížili ešte dvakrát. Grafické znázornenie sa stáva pri väčšom časovom období pomerne neprehľadné, ďalej sa teda bližšie pozrieme na porovnanie diagramu zaťaženia niektorých jednotlivých dní.



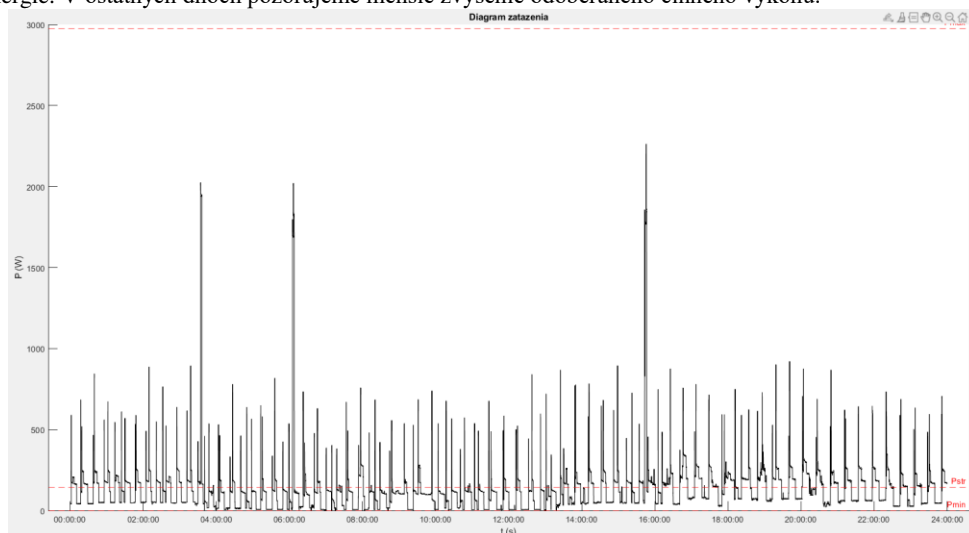
Obr. 3 Diagram zaťaženia – celé sledované obdobie



Obr. 4 Diagram zaťaženia zo dňa 14/04/2021

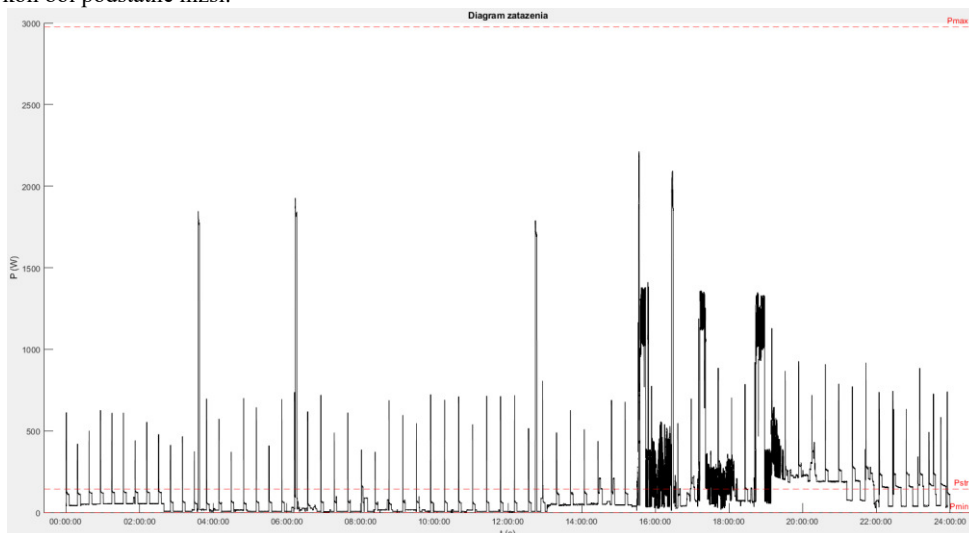
Typické príklady priebehu činného výkonu počas bežného pracovného dňa môžeme vidieť na obrázkoch 4 a 5. Na prvom z tejto dvojice obrázkov vidíme v popoludňajších hodinách a vo večernom čase mierne zvýšenie odoberaného výkonu. Takéto zvýšenie nastalo ešte počas ďalších troch dní. Jednalo sa o štvrtok 17/04/2021, pondelok 19/04/2021 a utorok 20/04/2021, teda bežné pracovné dni, kedy bolo v domácnosti počas dňa menej ľudí, ktorí aktívne využívali elektrickú energiu. Podobne, nárast odoberaného výkonu koreluje so západom slnka, teda nutnosťou využiť umelé elektrické osvetlenie.

Druhý typ diagramu zaťaženia, ktorý bol nameraný, môžeme vidieť na spomínanom obrázku 5. Takýto priebeh nastal v štyroch prípadoch, v nedeľu 16/04/2021, v utorok 18/04/2021, v piatok 21/04/2021 a v sobotu 22/04/2021. Pri porovnaní týchto priebehov si môžeme všimnúť, že na obrázku 4 nenastalo také výrazné zvýšenie odoberaného činného výkonu v poobedných a večerných hodinách ako na obrázku 6. Tento rozdiel pripisujeme do značnej miery množstvu osôb nachádzajúcich sa v domácnosti v daných dňoch. Tento počet sa menil priebežne v rozmedzí od 3 do 6. Práve v dňoch, kedy poobedné a večerné hodiny trávil v danej domácnosti viac osôb, odpovedajú dňom s viditeľne zvýšenou spotrebou elektrickej energie. V ostatných dňoch pozorujeme menšie zvýšenie odoberaného činného výkonu.



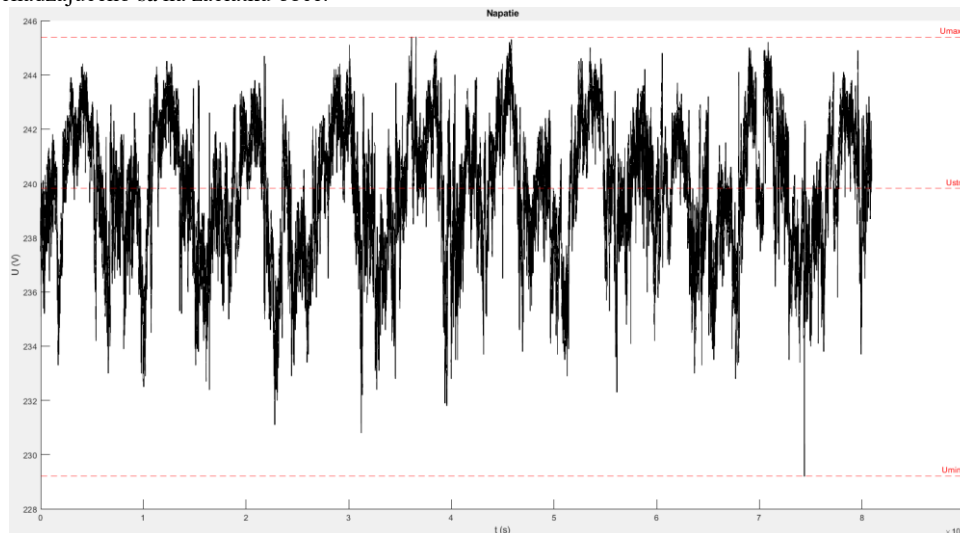
Obr. 5 Diagram zaťaženia zo dňa 22/04/2021

Ako samostatnú kategóriu uvedieme priebeh odoberaného činného výkonu domácnosti v sobotu 15/04/2021, ktorý môžeme vidieť na obrázku 6. Jedná sa o deň, v ktorom vidíme najväčší nárast spotreby elektrickej energie. Tento nárast je najviac viditeľný z dôvodu výskytu najväčšieho počtu osôb v domácnosti za sledované obdobie, ale rovnako aj kvôli činnostiam, ktoré v daný deň obyvatelia domácnosti vykonávali. Jedná sa o týždenné upratovanie, teda zvýšená záťaž na domovú elektroinštaláciu v podobe väčšieho využitia domácich spotrebičov, ako napríklad vysávač, elektrický parný čistič podláh, či práčka. Tieto faktory preukázateľne značne zvýšili spotrebu elektrickej energie, čo potvrdzuje aj porovnanie s nameranými údajmi z nasledujúcej soboty 22/04/2021, kde nenastal taký výrazný nárast spotreby elektrickej energie. V danú sobotu, obdobné upratovanie nenastalo, teda celkový odoberaný činný elektrický výkon bol podstatne nižší.



Obr. 6 Diagram zaťaženia zo dňa 15/04/2021

Z ďalších meraných veličín spomenieme napätie, ktorého priebeh počas celého sledovaného obdobia môžeme vidieť na obrázku 7. Z týchto nameraných údajov si môžeme všimnúť hlavne červenou vyznačené hodnoty, a to maximálne a minimálne namerané napätie a strednú hodnotu napätia. Prvé menované  $U_{\max}$  dosiahlo počas nášho merania hodnotu približne 245,5 V. Táto hodnota prekračuje hodnotu nominálneho napájacieho napätia bežných domových elektroinštalácií o 15 V. Priemerná hodnota napätia za celé sledované obdobie bola približne 240 V. Z tohto sa dá konštatovať dlhodobá nepriaznivá záťaž na spotrebiče, ktoré sú pravidelne využívané v danej domácnosti. Tento fakt potvrdzuje aj skutočnosť, že túto hodnotu nominálneho napätia sme namerali prakticky iba pri najnižšom dosiahnutom napätí, mierne pod 230 V. Tieto pracovné podmienky pre elektrospotrebiče v domácnosti sú zapríčinené geografickou polohou domu v rámci obce. Jedná sa o v poradí prvý dom od transformátora verejného rozvodu elektrickej energie nachádzajúceho sa na začiatku obce.



Obr. 7 Grafický priebeh napätia – celé sledované obdobie

## V. ZÁVER

V tomto článku sme sa venovali návrhu modelu domáceho elektromera, pomocou ktorého by bolo schopné merať, zaznamenávať a následne analyzovať základné elektrické veličiny v domovej elektroinštalácii. Vybrali sme vhodné súčiastky, predstavili princíp fungovania algoritmu zaznamenávania a spracovávania údajov a priblížili si prostredie, v ktorom má dané zariadenie pôsobiť, teda domovú elektroinštaláciu. V článku boli prezentované namerané výsledky, pričom sa jednalo o priebehy odoberaného činného výkonu počas niektorých vybraných dní merania a zhodnotenie faktorov, ktoré na tento odber vplývali. Rovnako boli v článku uvedené zistenia, ktoré boli pozorované našim zariadením a to je výška napätia v danej domácnosti, ktorá sa dlhodobo pohybovala výrazne nad nominálnou hodnotou napájacieho napätia 230 V, čo negatívne pôsobí na elektrické spotrebiče, aj samotnú elektroinštaláciu, ktoré sú využívané v tejto domácnosti.

## POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576, Kultúrnej a edukačnej grantovej agentúre MŠVVaŠ SR pod číslom KEGA 013TUKE-4/2021 a Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu SR a SAV podporenej grantom VEGA 1/0757/21.

## LITERATÚRA

- [1] *What is Arduino?* [online]. [cit. 2021-02-13]. Dostupné z: <https://www.arduino.cc/en/Guide/Introduction>
- [2] *Arduino Products* [online]. [cit. 2021-02-14]. Dostupné z: <https://www.arduino.cc/en/Main/Products>
- [3] *Peacefair Official Store*. [online]. [cit. 2021-02-14]. Dostupné z: <https://peacefair.aliexpress.com/store/1773456>
- [4] Selecký, M. *Arduino: uživatelská príručka*. Brno : Computer Press, 2016. ISBN 978-80-251-4840-2
- [5] Druga, C., Serban, I., Serbanescu, A., Cotoros, D. *Designing and Creating a Smart Audiometer Using an Arduino Platform*. [online]. 2021. Transylvania University, Brasov, Romania. [cit. 2021-02-21]
- [6] Zaplatílek K., Donar B. *MATLAB - tvorba uživatelských aplikácií*. Praha: BEN - technická literatúra, 2004. ISBN 80-7300-133-0
- [7] Varga, L., Ilenin, S. *Elektrické inštalácie a stanice*. Košice : Technická univerzita v Košiciach, 2012. ISBN 978-80-553-1176-0
- [8] Kanálik, M., Tomčík, J., *Prenos a rozvod elektrickej energie*. Košice : Technická univerzita v Košiciach, 2019. ISBN 978-80-553-3376-2
- [9] Hrivniak, P.: *Využitie open-source platformy pre elektroenergetické účely*. Bakalárska práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2021.

# Bezpečná vzdialená aktualizácia IoT senzora na báze ESP32

<sup>1</sup>Martin CHLEBOVEC, <sup>2</sup>Miloš DRUTAROVSKÝ

<sup>1,2</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>martin.chlebovec.2@student.tuke.sk, <sup>2</sup>milos.drutarovsky@tuke.sk

**Abstrakt** — Článok opisuje možnosti bezpečnej aktualizácie firmvéru IoT senzora na báze ESP32 v prostredí ESP-IDF, pričom kladie dôraz na prenos aktualizácie cez zabezpečený prenosový kanál, overenie integrity aktualizácie, zabezpečenie bootovacieho procesu a šifrovanie flash pamäte. Kombinácia týchto metód zaručuje vysokú mieru bezpečnosti pred najbežnejšími typmi útokov, ktoré sú cieleňé na IoT zariadenia.

**Kľúčové slová** — ESP32, OTA, digitálny podpis, ESP-IDF, zabezpečený bootovací proces, šifrovanie flash pamäte, firmvér

## I. ESP32 – ESPRESSIF SYSTEMS

ESP32 je mikrokontrolér z produkcie čínskej firmy Espressif Systems pre integráciu do IoT (internet vecí) aplikácií [1]. Podporuje WiFi (2,4 GHz) a Bluetooth konektivitu, pričom obe technológie zdieľajú spoločnú plošnú anténu na doske plošných spojov, čím sa výrazne redukuje aj samotná veľkosť modulu. WiFi technológia môže byť použitá pre prenos údajov z IoT senzora, ale aj na prevzatie OTA (Over-The-Air) aktualizácie firmvéru zo vzdialeného úložiska (servera).

Modul s ESP32 je vybavený externou flash pamäťou, najčastejšie s veľkosťou 4 MB. Flash pamäť slúži na uloženie nastavení WiFi adaptéra, firmvéru, softvérového bootloadera (zavádzača) a tabuľky partícií, ktorá rozdeľuje flash pamäť logickým členením na partície. ESP32 je najčastejšie vybavené dvojjadrovým procesorom harvardskej architektúry Tensilica Xtensa L6 s taktom až 240 MHz. Hlavný procesor Xtensa môže realizovať výpočty, komunikovať so senzormi a perifériami cez dostupné zbernice.

Zároveň však riadi a obsluhuje aj WiFi/Bluetooth zásobník (stack) pre zabezpečenie konektivity. Za správu a obsluhu WiFi/Bluetooth zásobníka zodpovedá jedno z jadier procesora. u ktorého má WiFi/Bluetooth zásobník maximálnu prioritu. Druhé jadro obsluhuje používateľskú aplikáciu. Obe jadrá zdieľajú spoločnú vyrovnávaciu pamäť.

## II. VLASTNOSTI A NÁSTROJE PROSTREDIA ESP-IDF

ESP-IDF je framework z produkcie Espressif Systems pre vývoj IoT aplikácií založený na jazyku C pre platformu ESP32 [2]. Je vyvíjaný na Githubu s poslednou produkčnou verziou 4.2. Framework ponúka kompatibilné knižnice a ukázkové projekty pre základné programové implementácie pre obsluhu zbernic, rozhraní ako základ pre vývoj vlastnej aplikácie na ESP32.

Umožňuje vyvíjať aplikácie a komplexnejšie programy s využitím plánovača a operačného systému reálneho času – FreeRTOS [3]. Výhodou plánovača je možnosť spúšťania vlastných programov na ESP32 ako samostatné úlohy, pričom je možné každej úlohe priradiť veľkosť zásobníka. V prípade pretečenia zásobníka sa daná úloha ukončí, avšak neovplyvní ostatné úlohy hlavného programu, ktoré sa naďalej vykonávajú.

Súčasťou frameworku sú aj vývojárske nástroje pre implementáciu bezpečnostných funkcionalít a taktiež Python skripty, ktoré dokážu vykonávať rôzne operácie. Skripty sú spúšťané v konzolovej aplikácii prostredia ESP-IDF s príslušným argumentom pre vykonanie konkrétnej úlohy. Základným Python skriptom v prostredí ESP-IDF je *idf.py*, ktorý dokáže obsluhovať príkazy súvisiace s kompiláciou programu, ale aj pre spustenie podprogramov prostredia. Umožňuje vyvolať Menuconfig (konfiguračné menu), ktorým je možné nastaviť aktuálne otvorený projekt pre potreby vývojára. Dokáže vykonať zápis skompilovaného firmvéru a ďalších skompilovaných obrazov do flash pamäte.

Externá flash pamäť môže byť rozdelená tabuľkou partícií, ktorá sa kompiluje spoločne

s hlavnou aplikáciou na logické partície, ktoré môžu byť aplikačné (s podporou), systémové a taktiež príznakové, ak sa používa viac aplikačných partícií pre nastavenie príznaku pre bootovanie (zavedenie) firmvéru z preferovanej partície.

Pre vzdialenú aktualizáciu firmvéru sa využíva rozdelenie partícií schémou „Factory App, two OTA definitions“, ktorá umožňuje uložiť a bootovať až 3 firmvéry z flash pamäte, pričom partícia Factory je vždy vyhradená pre firmvér nahrať cez fyzické USB-UART rozhranie a slúži zároveň ako „fail-safe“ riešenie v prípade, že sa nepodarí bootovať firmvér z iných – OTA partícií (OTA\_0 a OTA\_1), ktoré sú použité pre uloženie aktualizovaného firmvéru.

Ďalším z dostupných Python skriptov v prostredí ESP-IDF je *esptool.py*, ktorý slúži pre prácu s flash pamäťou ESP32. Dokáže vykonať priamy zápis do flash pamäte (najčastejšie je spúšťaný skriptom *idf.py*).

Označenie „ESPTOOL“ [4] je zároveň aj označením pre balík Python skriptov dostupných v prostredí ESP-IDF (*esptool.py*, *espsecure.py* a *espefuse.py*). Pre kryptografické operácie je v prostredí ESP-IDF dostupný skript *espsecure.py*. Operácie súvisiace s jednorazovo programovateľnými pamäťami eFuses (4x 256-bit) obsluhuje skript *espefuse.py*.

### III. VZDIALENÁ AKTUALIZÁCIA FIRMVÉRU V PROSTREDÍ ESP-IDF

ESP-IDF obsahuje viacero ukázkových projektov pre vzdialenú aktualizáciu firmvéru. Projekt Native OTA [5] umožňuje vykonať vzdialenú aktualizáciu firmvéru z webservera, ktorý aktualizáciu distribuuje s využitím bezpečného HTTPS (Hypertext Transfer Protocol Secure) protokolu s end-to-end šifrovaním pre architektúru klient-server. Na webserver vykonáva mikrokontrolér ESP32 (klient) jednorazovú GET požiadavku pre prevzatie obsahu firmvéru (.bin obraz). Aktualizovaný firmvér je zapísaný do partície OTA\_0, alebo OTA\_1, odkiaľ je následne bootovaný.

Pre realizáciu zabezpečeného spojenia na strane ESP32 sa vyžaduje certifikát certifikačnej autority v .pem formáte, ktorý sa vloží do preddefinovaného priečinka projektu, kde ho očakáva kompilátor a linker. Certifikačná autorita vydala certifikát pre webserver, na ktorý realizuje pripojenie ESP32. Webserver predkladá klientovi certifikát pri realizácii zabezpečeného spojenia. ESP32 dôveruje certifikátu webservera, keďže ho vydala dôveryhodná certifikačná autorita, ktorej certifikátom klient disponuje. Využíva sa Chain of Trust (reťaz dôveryhodnosti).

Projekt Native OTA sa konfiguruje cez Menuconfig prostredia ESP-IDF. Konfigurovať je možné cestu k cieľovému firmvéru (URL), kde ho distribuuje webserver a taktiež aj údaje súvisiace s konektivitou (SSID, heslo WiFi siete). Native OTA umožňujú vykonávať aj kontrolu verzie firmvéru. Dokáže porovnať verziu aktuálne bežiaceho a stiahnutého firmvéru. V prípade, že sú verzie rozdielne, prepíše sa príznak pre primárne bootovanú partíciu v OTA\_DATA partícií a vykoná softvérový reštart pre bootovanie aktualizovaného firmvéru. V prípade totožných verzií sa reštart, ani zmena v OTA\_DATA partícií nevykoná.

### IV. ZABEZPEČENIE AKTUALIZÁCIE FIRMVÉRU – DIGITÁLNY PODPIS

Keďže projekt Native OTA nemá mechanizmy pre overenie integrity firmvéru a bootloadera, je nutné tieto bezpečnostné funkcionality implementovať dodatočne prostredníctvom nástrojov nato určených, ktoré sú obsiahnuté v Menuconfigu prostredia ESP-IDF. Výhodou funkcionality je, že ich nie je nutné implementovať do zdrojových kódov aplikácie.

Metóda digitálneho podpisu [6] umožňuje overiť integritu firmvéru. Overený firmvér je bezpečný pre spustenie na vývojovej platforme ESP32. Digitálny podpis vychádza z asymetrickej kryptografie, ktorá používa dvojicu kľúčov – súkromný a verejný. Súkromný kľúč je použitý pre manuálne podpísanie firmvéru vývojárom. Po podpísaní firmvéru ho môže vývojár sprístupniť na OTA serveri, odkiaľ je distribuovaný klientom, ktorí ho sú schopní overiť v prípade, že majú verejný kľúč podpisujúceho k dispozícii.

Verejný kľúč je v procese kompilácie vložený do obrazu softvérového bootloadera. Softvérový bootloader realizuje overenie firmvéru pri jeho bootovaní s možnosťou overenia firmvéru aj v procese samotnej aktualizácie „ON UPDATE“.

Prostredie ESP-IDF využíva kryptografiu na báze ECC (Elliptic Curve Cryptography), ktorá je založená na zložitosti hľadania diskretného logaritmu pre náhodný bod eliptickej krivky nad konečným poľom. Pre samotný digitálny podpis sa využíva podpisová schéma ECDSA (Elliptic Curve Digital Signature Algorithm).

Pre implementáciu digitálneho podpisu v prostredí ESP-IDF je možné použiť skript *espsecure.py*, ktorý je možné spustiť v prostredí ESP-IDF cez konzolovú aplikáciu frameworku. Príkazom *espsecure.py generate\_signing\_key private.pem* je možné vygenerovať súkromný kľúč pre eliptickú krivku „NIST256p“. Generovanie sa realizuje s využitím entropie operačného systému, na ktorom je ESP-IDF spustené. Ako alternatívu je možné použiť aj kryptografický

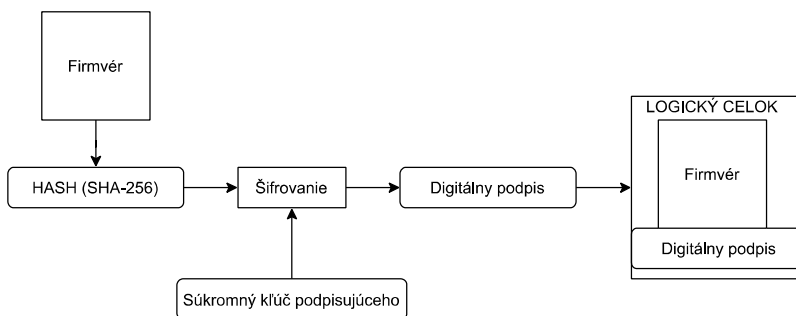
nástroj OpenSSL, ktorý umožňuje generovať kľúč pre identickú eliptickú krivku s označením „prime256v1“. Výstupom skriptu je súkromný kľúč „private.pem“ s dĺžkou 256 bitov v .pem formáte.

Verejný kľúč je možné skriptom *espsecure.py* vygenerovať zo súkromného kľúča príkazom *espsecure.py extract\_public\_key --keyfile private.pem public.bin*. Aby bolo možné metódu digitálneho podpisu pre aplikáciu založenú na ESP32 použiť, je potrebné v prostredí ESP-IDF v Menuconfigu v časti „Security Features“ zapnúť možnosti „Require signed apps“ a „Verify via Bootloader on startup“, pričom je možné digitálny podpis overiť aj priamo v procese stiahnutia firmvéru, pri akcii „ON\_UPDATE“, ktorú je v menu tiež možné vybrať.

Ďalším parametrom, ktorý je potrebné nastaviť je relatívna cesta (vzhľadom na koreňový priečinok projektu) k verejnému kľúču – „public.bin“, ktorý sa v procese kompilácie vloží do obrazu softvérového bootloadera. Podpísanie firmvéru je potrebné realizovať manuálne po kompilácii firmvéru a pred jeho nahrávaním do ESP32 cez USB-UART rozhranie, alebo pred vložení firmvéru na OTA webserver.

Pre podpísanie firmvéru je možné použiť *espsecure.py* a príkazom *espsecure.py sign\_data --version 1 --keyfile private.pem --output native\_ota.bin native\_ota.bin* sa vykoná podpísanie firmvéru „native\_ota.bin“ súkromným kľúčom „private.pem“.

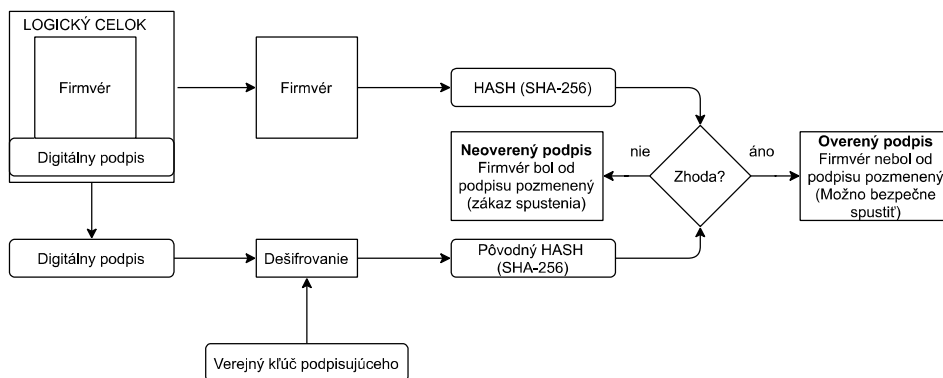
V procese podpisovania firmvéru sa z pôvodného súboru „native\_ota.bin“ vypočíta hašovaná hodnota funkciou SHA256 (Secure Hash Algorithm). Odtlačok je zašifrovaný súkromným kľúčom a jeho výsledkom je digitálny podpis, ktorý je vložený do firmvéru za jeho pôvodný obsah. Podpis má dĺžku 68 bajtov, pričom prvé 4 bajty sú slovo tvorené nulami a nasledujúcich 64 bajtov je samotný digitálny podpis. Bloková schéma podpisania firmvéru vývojárom je opísaná na Obr. 1.



Obr. 1 Proces podpisania firmvéru súkromným kľúčom

Verejný kľúč podpisujúceho je zapísaný v obraze softvérového bootloadera. Bootloader kľúčom dešifruje digitálny podpis a získa pôvodnú hašovanú hodnotu SHA256 firmvéru. Následne vytvorí hašovanú hodnotu SHA256 firmvéru a porovná obe hašované hodnoty. V prípade, že sa zhodujú, firmvér je autentický a nebol pozmenený treťou osobou od momentu podpisania.

Firmvér je dôveryhodný a je ho možné spustiť bezpečne na ESP32. Úspešnosť samotného overenia spustí následne funkciu z programu Native OTA, ktorá overuje, či je verzia aktuálne spusteného firmvéru iná, ako verzia stiahnutého firmvéru. V prípade rozdielnych verzií firmvérov sa vykoná prepis v príznakovej partícii OTA\_DATA, ktorá definuje primárne bootovanú partíciu s firmvérom a vykoná sa softvérový reštart pre možnosť bootovania aktualizovaného firmvéru z flash pamäte. Proces overenia firmvéru je blokovo schémou opísaný na Obr. 2.



Obr. 2 Proces overenia firmvéru verejným kľúčom

## V. ZABEZPEČENIE BOOTOVACIEHO PROCESU – SECURE BOOT

Secure Boot (SB) je metóda zabezpečenia bootovacieho procesu. Kombinuje použitie metódy digitálneho podpisu firmvéru a jeho overenia softvérovým bootloaderom v procese bootovania a aktualizácie pre akciu „ON\_UPDATE“. Zároveň je SB [6] rozšírená o overenie softvérového bootloadera na začiatku bootovacieho procesu, ktorú obsluhuje hardvérový bootloader uložený v ROM pamäti.

Pre produkčné aplikácie je vhodné použiť SB s „One-Time Flash“ nastavením, ktoré využíva AES (Advanced Encryption Standard) symetrický kľúč uložený v jednorazovo programovateľnej pamäti eFuse, odkiaľ ho dokáže prečítať a použiť výhradne hardvérový bootloader, keďže eFuse je chránená pred softvérovým prečítaním jej obsahu, či prepisu jej obsahu. V princípe ide o výpočet odtlačku obrazu softvérového bootloadera uloženého vo flash pamäti na ofsete 0x1000 algoritmom SBDA (Secure Bootloader Digest Algorithm).

Vypočítaný odtlačok je následne porovnaný s referenčným, ktorý je uložený vo flash pamäti na preddefinovanom ofsete 0x0, kde ho zapisuje vývojár, ktorý ho vygeneroval lokálne. Výpočet odtlačku a porovnanie na strane ESP32 realizuje hardvérový bootloader. V prípade, že je odtlačok identický, softvérový bootloader je dôveryhodný. Hardvérový bootloader umožní softvérovému bootloaderu prístup k bootovaniu firmvéru a overenia jeho digitálneho podpisu. V prípade, že sú odtlačky rozdielne, fáza bootovania firmvéru je zakázaná a ESP32 sa reštartuje v nekonečnej slučke a skúša opätovné overenia obrazu softvérového bootloadera.

Algoritmus SBDA využíva AES kľúč (256-bit) z eFuse BLK2, ktorý načíta v reverznej bitovej reprezentácii a obraz softvérového bootloadera z flash pamäte z ofsetu 0x1000. Vygeneruje sa 128-bajtový inicializačný vektor, ktorý sa dosadí pred bootloader. Vykoná sa zarovnanie obrazu bootloadera modulo 128 a následne jeho dorovnanie pre modulo 128 s doplnením hodnôt 0xFF do reprezentácie obrazu.

Na každých 16 bajtov otvoreného textu obrazu bootloadera sa aplikuje bloková šifra AES256 v ECB móde s využitím AES kľúča z eFuse BLK2 s dĺžkou 256 bitov. Výsledný šifrovaný text má reverznú bitovú reprezentáciu. Algoritmus vymení bajt každého 4-bajtového slova šifrovaného textu, vypočíta hašovanú hodnotu SHA512 výsledného šifrovaného textu. Výstupom je 192-bajtový reťazec, ktorý je tvorený 128-bajtovým inicializačným vektorom a 64-bajtovou hašovanou hodnotou SHA512 zo šifrovaného textu.

Vývojár, ktorý lokálne generuje odtlačok má k dispozícii AES kľúč, ktorý predtým vygeneroval a je identický s kľúčom zapísaným v bloku eFuse BLK2. Príkazom *espsecure.py generate\_generate\_flash\_encryption\_key\_secure-bootloader-key-256.bin* je možné vygenerovať AES kľúč „secure-bootloader-key-256.bin“. Pre generovanie sa používa Python funkcia *os.random()*, ktorá využíva viac faktorov operačného systému, na ktorom je spustené rozhranie ESP-IDF pre vyššiu entropiu – náhodnosť kľúča.

SB je možné zapnúť cez Menuconfig, v časti „Security Features“, avšak pre permanentné zapnutie je nutné zapísať potvrdzovací bit do 1-bitovej eFuse ABS\_DONE\_0. Príkazom *espefuse.py burn\_key\_secure\_boot\_secure-bootloader-key-256.bin* sa vykoná zápis AES kľúča do eFuse BLK2. Potvrdzovací bit do eFuse ABS\_DONE\_0 je možné zapísať príkazom *espefuse.py burn\_efuse\_ABS\_DONE\_0*. Potencionálny útočník nedokáže získať kľúč uložený v tejto eFuse, keďže k nej nemá prístup žiadnym softvérovým nástrojom.

K bloku eFuse BLK2 môže pristupovať iba hardvérová funkcionálna SB prostredníctvom hardvérového bootloadera uloženého v ROM pamäti. Táto metóda je efektívna pred spustením modifikovaného bootloadera, ktorý by mohol útočník do ESP32 nahrat cez fyzické USB-UART rozhranie, čím by mohol spustiť bootovanie nepodpísaného (nebezpečného) firmvéru, ktorý by mohol byť podvrhnutý v aktualizácii, alebo nahratý cez fyzické USB-UART rozhranie do ESP32.

## VI. ŠIFROVANIE FLASH PAMÄTE – FLASH ENCRYPTION

Flash Encryption (FE) je metóda šifrovania flash pamäte. Umožňuje šifrovať istú časť flash pamäte, alebo celý jej obsah [7]. Každá šifrovaná partícia je v tabuľke partícií definovaná príznakom encrypted. Šifrovaná časť flash pamäte nie je potencionálnym útočníkom spustiteľná (napríklad firmvér), keďže dokáže prevziať cez USB-UART rozhranie iba zašifrovaný firmvér, ktorý nie je bez symetrického AES kľúča spustiteľný, keďže musí byť dešifrovaný.

V otvorenom texte môže byť naďalej distribuovaná aktualizácia firmvéru, ktorý je zašifrovaný v procese zápisu do flash pamäte do aplikačnej partície označenej príznakom encrypted. FE používa symetrický AES kľúč s dĺžkou 256 bitov pre šifrovanie a dešifrovanie obsahu flash pamäte. Kľúč je uložený v jednorazovo programovateľnej pamäti eFuse BLK1 s veľkosťou 256 bitov, ktorá je pre tento kľúč vyhradená. K pamäti eFuse BLK1 dokáže pristupovať iba hardvérový bootloader, ktorý operácie šifrovania a dešifrovania vykonáva.



Všetky aplikačné particie, tabuľka particií, softvérový bootloader a referenčný odtlačok je vždy šifrovaný bez ohľadu na príznak. Funkcionalitu FE je možné zapnúť cez Menuconfig v časti „Security Features“. Režim „Release“ je vhodný pre produkčné aplikácie, keďže šifrovanie flash pamäte nie je možné v budúcnosti vypnúť z dôvodu, že je príznaková eFuse stavu šifrovania flash pamäte chránená proti prepisu.

Po skompilovaní bootloadera s nastaveným príznakom pre funkcionalitu FE a jeho nahrati do ESP32 je obsah flash pamäte nešifrovaný. Za predpokladu, že je spustená aj funkcionlita SB, hardvérový bootloader overí softvérový bootloader pre možnosť spustenia ďalších fáz bootovacieho procesu. Po úspešnom overení softvérový bootloader na základe príznaku šifrovania flash pamäte načíta hodnotu bloku eFuse FLASH\_CRYPT\_CNT.

Ak je jej hodnota 0 (ešte nešifrovaná flash pamäť), nastaví a aktivuje sa blok šifrovania obsahu flash pamäte, bootloader nastaví 4-bitovú eFuse FLASH\_CRYPT\_CONFIG na hodnotu 0xF. Operácie samotného šifrovania flash pamäte už vykonáva hardvérový bootloader, keďže softvér nemá prístup k bloku eFuse BLK1, kde je uložený symetrický AES kľúč.

Pre šifrovanie aj dešifrovanie sa využíva AES kľúč zapísaný do eFuse BLK1 s veľkosťou 256-bitov, ktorá je na tento účel určená. Pri väčších particiách môže trvať šifrovanie až minútu. Softvérový bootloader po ukončení šifrovania nastaví eFuse FLASH\_CRYPT\_CNT na hodnotu 0x01, čo znamená, že je obsah flash pamäte šifrovaný a opätovné šifrovanie sa pri reštarte a spúšťaní systému nevykoná.

Algoritmus šifrovania flash pamäte využíva blokovú šifru AES256, ktorá pracuje na 32-bajtových blokoch otvoreného textu, využíva dva bloky AES v sérii. AES256 používa symetrický AES v reverznej bitovej reprezentácii. Pre proces šifrovania a dešifrovania flash pamäte sa používajú kryptografické operácie AES opačne. Proces šifrovania je tvorený funkciou AES decrypt a proces dešifrovania flash pamäte funkciou AES encrypt.

Pri šifrovaní sa na každý 32 bajtový blok otvoreného textu aplikuje unikátny AES kľúč, ktorý je odvodený od hlavného AES kľúča vykonaním operácie XOR (exkluzívny súčet) medzi AES kľúčom z eFuse BLK1 a ofsetom bloku otvoreného textu vo flash pamäti. XOR-ovanie konkrétnych bitov kľúča závisí od hodnoty v 4-bitovej eFuse FLASH\_CRYPT\_CONFIG.

To zaručuje, že sú XOR-ované všetky bity AES kľúča s ofsetom bloku dát. Zmena hodnoty zapísanej v eFuse FLASH\_CRYPT\_CONFIG by mohla znížiť kryptografickú bezpečnosť šifrovania flash pamäte, keďže by sa operácia XOR nevykonala pre všetky bity AES kľúča. Pre 1. nastavený bit sa vykonáva XOR pre bity 0 až 66, pre 2. bit hodnoty 67 až 131, pre 3. bit hodnoty 132 až 194, pre 4. bit hodnoty 195 až 256.

Prvým krokom pre spustenie FE je vygenerovanie a zapísanie symetrického kľúča do príslušnej jednorazovo programovateľnej pamäte eFuse BLK1. S využitím nástroja *espsecure.py* a príkazom *espsecure.py generate\_flash\_encryption\_key my\_flash\_encryption\_key.bin* je možné kľúč vygenerovať. Funkcia pre generovanie AES kľúča je totožná ako v prípade kľúča pre SB, ktorá vychádza z náhodnosti operačného systému.

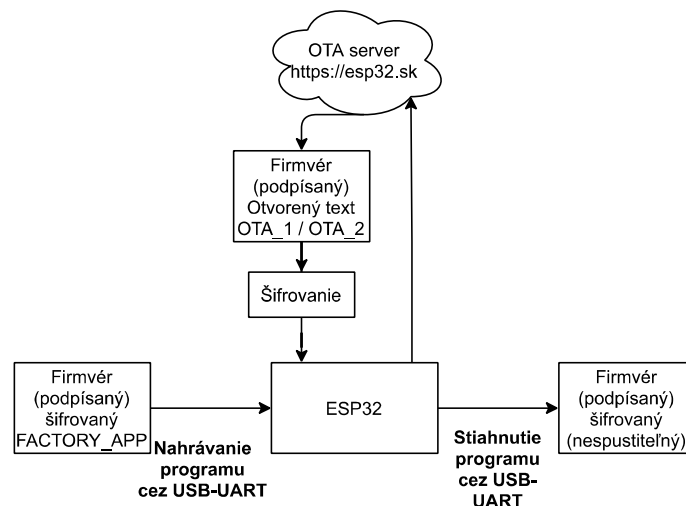
Vygenerovaný kľúč „my\_flash\_encryption\_key.bin“ je následne potrebné zapísať do vyhradenej jednorazovo programovateľnej pamäte eFuse BLK1, ktoré je pre neho určená. Príkazom *espefuse.py -port PORT burn\_key flash\_encryption my\_flash\_encryption\_key.bin* sa kľúč zapíše do eFuse BLK1.

Šifrovaný firmvér je možné distribuovať vo vzdialenej aktualizácii v prípade úpravy projektu Native OTA pre zmenu zápisu funkcie firmvéru do flash pamäte na *esp\_partition\_write()*, ktorá však vyžaduje nastavenie začiatočného ofsetu particie, kde bude firmvér zapísaný. Tento ofset bude použitý aj pre operáciu XOR na ESP32 medzi ofsetom dát a AES kľúčom z pamäte eFuse pre výpočet odvodeného kľúča, ktorým sa následne blok dát šifruje. Z toho dôvodu je možné využívať iba jednu OTA partíciu pre budúce aktualizácie, keďže inak by firmvér nebol dešifrovaný do spustiteľnej podoby správne a zlyhalo by tak jeho bootovanie.

Pre šifrovanie obrazov (firmvér, odtlačok, softvérový bootloader) sa používa príkaz *espsecure.py encrypt\_flash\_data*, ktorý má viacero parametrov, napr. začiatočný ofset pre správne vykonanie operácie XOR. Python skript počíta s nastavením eFuse FLASH\_CRYPT\_CONFIG na štandardnú hodnotu 0xF pre vykonanie operácie XOR pre všetky bity AES kľúča, ktorý je tiež jedným z parametrov – „my\_flash\_encryption\_key.bin“.

Všetky opísané metódy boli použité v diplomovej práci „Bezpečná aktualizácia firmvéru v senzorovej sieti na báze ESP32“ [8] pre demonštráciu senzorového uzla, ktorý okrem zberu údajov z meteorologického senzora a ovládania digitálneho výstupu na spôsob izbového termostatu vykonáva bezpečnú aktualizáciu firmvéru zo vzdialeného webservera.

Na Obr. 3 je vizualizácia FE pre proces vzdialenej aktualizácie firmvéru, nahratie firmvéru cez fyzické USB-UART rozhranie a pre prevzatie firmvéru, resp. obsahu flash pamäte s príznakom encrypted.



Obr. 3 Metóda šifrovania flash pamäte pre stiahnutie, nahratie firmvéru, prevzatie vzdialenej aktualizácie – Použitie v DP pre finálnu aplikáciu pre režim Release [8]

### ZÁVER

Opísané metódy zabezpečenia firmvéru, bootovacieho procesu a šifrovania flash pamäte poskytujú vysokú úroveň bezpečnosti pred bežnými typmi útokov cielených na IoT zariadenia. Potencionálny útočník nedokáže na produkčnom mikrokontroléri ESP32 spustiť podvrhnutý bootloader, prípadne firmvér cez fyzické USB-UART rozhranie, alebo cez vzdialenú aktualizáciu. V prípade prevzatia firmvéru z flash pamäte dokáže útočník prevziať iba šifrovaný firmvér, ktorý nie je bez symetrického AES kľúča spustiteľný.

Kľúče k metóde zabezpečeného bootovacieho procesu a šifrovania flash pamäte sú bezpečne uložené v pamätiach eFuses, kde nemá softvér prístup. Zálohu kľúčov má k dispozícii vývojár, ktorý ich používa pre šifrovanie firmvéru a lokálny výpočet referenčného odtlačku bootloadera v prípade jeho úpravy. Útočník tak nedokáže kľúče získať a použiť ich, napríklad pri podvrhnutí firmvéru.

Súčasťou diplomovej práce [8] je detailnejší opis jednotlivých metód zabezpečenia aj s návodom na ich implementáciu v prostredí ESP-IDF pre produkčnú verziu frameworku 4.2. V prílohách diplomovej práce sú zdrojové kódy hlavnej aplikácie na ESP32 a webservera pre demonštráciu IoT senzorového uzla s podporou bezpečnej vzdialenej aktualizácie firmvéru s využitím vlastného webservera pre zber údajov zo senzorového uzla a pre distribúciu vzdialených (OTA) aktualizácií.

### POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-18-0373.

### REFERENCIE

- [1] ESP32 Technical Reference Manual [online]. Espressif Systems [cit. 2020-01-01]. Dostupné z: [https://www.espressif.com/sites/default/files/documentation/esp32\\_technical\\_referenc\\_e\\_manual\\_en.pdf](https://www.espressif.com/sites/default/files/documentation/esp32_technical_referenc_e_manual_en.pdf)
- [2] IoT Development Framework - Overview [online]. Espressif Systems [cit. 2021-01-20]. Dostupné z: <https://www.espressif.com/en/products/sdks/esp-idf>
- [3] The FreeRTOS™ Reference Manual [online]. Amazon.com, Inc. [cit. 2018-07-01]. Dostupné z: [https://www.freertos.org/fr-contentsrc/uploads/2018/07/FreeRTOS\\_Reference\\_Manual\\_V10.0.0.pdf](https://www.freertos.org/fr-contentsrc/uploads/2018/07/FreeRTOS_Reference_Manual_V10.0.0.pdf)
- [4] ESPTOOL (bundle) [online]. Github [cit. 2021-02-22]. Dostupné z: <https://github.com/espressif/esptool/>
- [5] Native OTA [online]. Github [cit. 2021-01-15]. Dostupné z: [https://github.com/espressif/esp-idf/blob/master/examples/system/ota/native\\_ota\\_example/main/native\\_ota\\_example.c](https://github.com/espressif/esp-idf/blob/master/examples/system/ota/native_ota_example/main/native_ota_example.c)
- [6] Secure Boot [online]. ESP-IDF Programming Guide [cit. 2021-02-27]. Dostupné z: <https://docs.espressif.com/projects/esp-idf/en/v4.2/esp32/security/secure-bootv1.html>
- [7] Flash Encryption [online]. ESP-IDF Programming Guide [cit. 2021-02-27]. Dostupné z: <https://docs.espressif.com/projects/esp-idf/en/latest/esp32/security/flashencryption.html>
- [8] CHLEBOVEC, M., Bezpečná aktualizácia firmvéru v senzorovej sieti na báze ESP32, Diplomová práca, Košice, Máj 2021, 97 s.

# Tepelná analýza priameho kontaktu

*Ján PRESADA, Dušan MEDVEĎ*

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

jan.presada@student.tuke.sk, dusan.medved@tuke.sk

**Abstrakt** — Tento článok sa venuje matematickému modelovaniu rozloženia teploty v okolí priameho elektrického vysokoprúdového kontaktu pri pôsobení nominálnej hodnoty prúdu 3000 A. Priame elektrické kontakty patria medzi prvky, pomocou ktorých sa pripája veľké množstvo elektrických zariadení. Dôležitú úlohu majú hlavne v prenosovej a distribučnej sústave, kde musia odolávať nepriaznivým poveternostným podmienkam, ktoré majú značný vplyv na ich degradáciu.

**KLúčové slová** — teplo, teplota, elektrický prúd, elektrický kontakt, rezistivita, koeficient tepelnej vodivosti

## I. ÚVOD

Elektrické kontakty zohrávajú v elektrických systémoch zdanlivo jednoduchú, ale predsa dôležitú úlohu. Predstavujú rozhranie medzi segmentmi obvodu, umožňujú segmentom pripojiť sa a odpojiť sa v prípade potreby. Napriek tomu, že to na prvý pohľad znie jednoducho, návrh tohto typu spojenia (elektrického kontaktu) si vyžaduje dôkladné naplánovanie a zváženie všetkých možných nepriaznivých vplyvov. Pri nesprávnom navrhovaní môžu elektrické kontakty vykazovať nadmerné zafarbenie, opotrebenie zvaráním, mechanické poškodenie alebo korozívne opotrebenie, čo môže mať za následok chybné pripojenie. Kontakty musia byť navrhnuté tak, aby sa minimalizoval výskyt týchto poruchových režimov a to všetko pri zachovaní nízkeho kontaktného odporu, minimálneho elektrického šumu a spoľahlivej schopnosti spojenia a odpojenia.

Prevažná časť kontaktov sa pri vypínaní opaľuje pôsobením elektrických oblúkov, ktoré vznikajú na kontaktoch pri prerušení elektrického obvodu. Elektrické kontakty preto patria medzi najporuchovejšie časti prístrojov a obvodov, preto sa na ich výrobu používajú špeciálne materiály. Existujú stovky rôznych kontaktných materiálov, ktoré sa úspešne používajú v mnohých rôznych aplikáciách.

## II. MATERIÁLY ELEKTRICKÝCH KONTAKTOV

V elektrotechnike sa používa veľké množstvo materiálov, ktoré sa vyrábajú v hutníckom a chemickom priemysle. Elektrotechnické materiály sú oveľa drahšie ako bežné materiály, pretože majú podstatne väčšiu čistotu. Preto je veľmi dôležité správne hospodárenie s nimi a ich následná recyklácia.

Rozoznávame 5 druhov elektrotechnických materiálov: elektricky vodivé (vodiče), elektricky nevodivé (izolanty), polovodivé (polovodiče), magnetické a konštrukčné materiály.

Elektricky vodivé materiály, používané v elektrotechnike, sú väčšinou pevné (kovy alebo nekovy), ale elektrický prúd vedú aj kvapaliny a plyny. V pevných vodičoch sa o vedenie prúdu starajú voľné elektróny a v kvapalinách a plyných ióny [1].

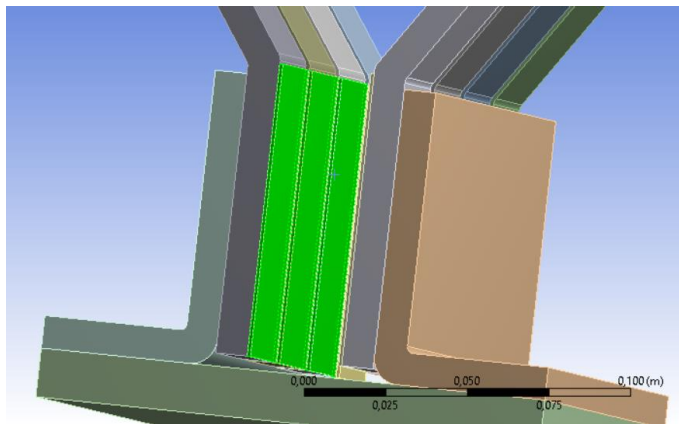
Na elektrické kontakty sú kladené požiadavky elektrického, mechanického a tepelného charakteru. Žiadny kontaktný styk nedokáže prenášať elektrickú energiu bez strát. Preto aj v elektrických kontaktoch vznikajú často nežiadúce javy, ktoré sa generujú prechodom prúdu, ako napríklad oteplenie materiálu, prenos materiálu a oxidácia kontaktov. Tieto javy sú, spolu s prechodovým odporom, veľmi dôležité pri voľbe materiálu pre elektrický kontakt.

Z hľadiska materiálov, do najnáročnejších skupín kontaktov patria spínacie prúdové obvody, spínané počas prevádzky. Materiálmi pre ich výrobu sú rôzne kovy a zliatiny, ktorých hlavnou vlastnosťou je tuhosť alebo pružná tvrdosť. V mnohých prípadoch sa na výrobu kontaktov používajú kontaktné dvojkovy, ktoré sa vyrábajú vo forme drôtov, pásov a profilov [2].

## III. VPLYV NEČISTÔT PRIAMEHO KONTAKTU NA ROZLOŽENIE TEPLoty

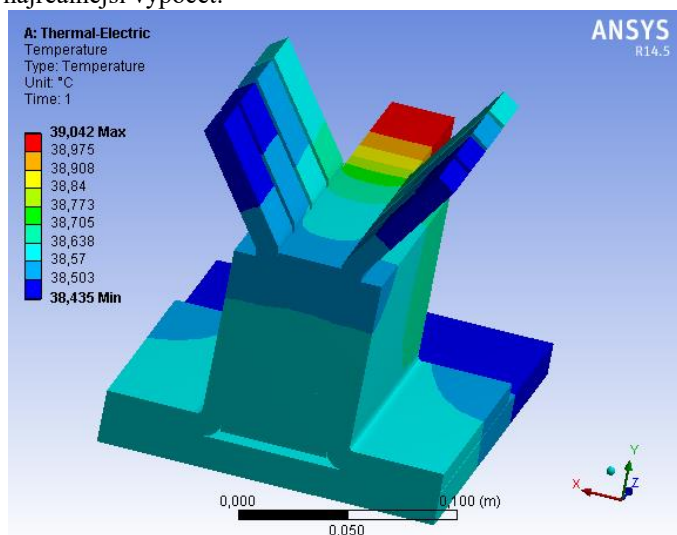
Vplyv nečistôt na teplotné rozloženie bolo zisťované tak, že sme na všetky kontakty pridali

plochu, ktorej šírka mala 1 mm. Táto plocha charakterizuje mierne opotrebovanie, nerovnosti a hlavne oxidáciu kontaktov, ktorá nastáva v dôsledku negatívneho vplyvu prostredia, starnutím materiálu a ďalších faktorov. V skutočnosti sú hrúbky takýchto nečistôt oveľa menšie, rádovo 10÷4 nm, ale kvôli zjednodušeniu výpočtu v prostredí ANSYS, by pri takýchto rozmeroch mohlo dôjsť k chybám pri vytváraní sieťovania na nečistotách. Na (Obr. 1) sú nečistoty zvýraznené zelenou farbou.



Obr. 1 Znáznorenie nečistôt na kontaktoch (zelenou farbou)

Med', ako materiál využívaný výrobcom, som pri výrobe daného druhu kontaktu použil na počiatočné simulácie. Na základe ktorých som zisťoval dôležitosť zadávaných parametrov a hodnôt, pre čo najreálnejší výpočet.



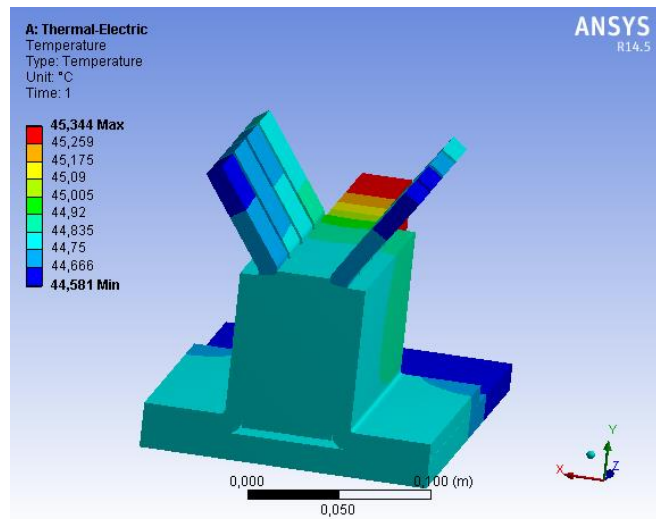
Obr. 2 Teplotné rozloženie priameho medeného kontaktu (bez nečistôt)

Z teplotného rozloženia na Obr. 2 je vidieť, že najvyššia teplota sa nachádza na začiatku noža kontaktu, s maximálnou hodnotou 39,042°C. Najnižšie teploty sú na tej časti podstavca, nad ktorou sa držiaky kontaktov nenachádzajú a zároveň na koncoch kontaktov uložených na konci noža, kde minimálna hodnota teploty dosahuje 38,435°C.

Model kontaktu na Obr. 2 bol namodelovaný bez spomínaných nečistôt, čiže zobrazuje tzv. „ideálny“ prevádzkový stav, počas ktorého nedochádza k prehrievaniu kontaktov, zapríčinené nečistotami. V tomto prípade by k prehriatiu mohlo dôjsť napríklad pri zvýšenej hodnote elektrického prúdu, ktorý preteká priamym kontaktom.

Pomocou uvedeného rozsahu rezistivity bola následne upravovaná zhoršujúca elektrická vodivosť, ktorá predstavovala znečistenie medzi kontaktmi a nožom. Pričom hodnota  $10^{-8} \Omega \cdot m$  predstavovala drobné nečistoty a hodnota  $10^{-4} \Omega \cdot m$  charakterizovala veľmi veľké znečistenie plôch.

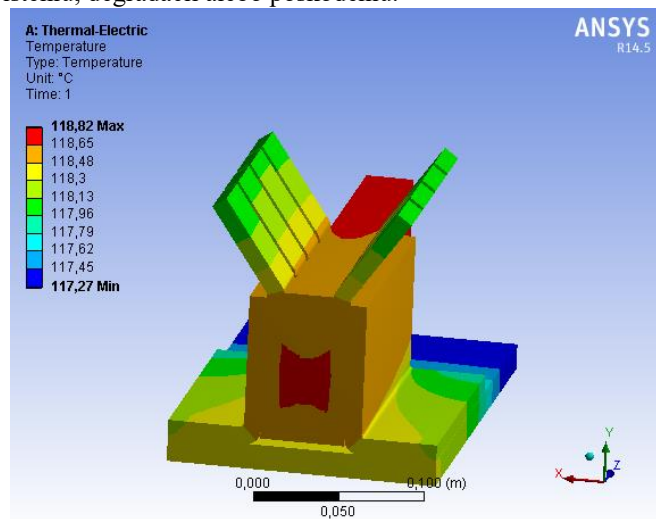
Na Obr. 3 je znázornený kontakt s rezistivitou  $1 \cdot 10^{-8} \Omega \cdot m$ . Uvažovalo sa, že táto hodnota vyjadruje „reálny“ prevádzkový stav pri počiatočnom uvedení do prevádzky.



Obr. 3 Priebeh rozloženia teplôt priameho medeného kontaktu, s rezistivitou nečistoty  $1 \cdot 10^{-8} \Omega \cdot m$

Dané rozloženie teplôt potvrdzuje vyššiu maximálnu aj minimálnu hodnotu teplôt oproti predošlému stavu (bez nečistoty), pričom došlo k zvýšeniu približne o  $6^\circ\text{C}$ . Najvyššie a najnižšie teploty sa nachádzajú na rovnakých miestach ako na Obr. 2, s tým rozdielom, že došlo k miernemu zväčšeniu niektorých pásiem. Tieto skutočnosti vyjadrujú odlišnosť medzi ideálnym a reálnym stavom priameho kontaktu.

Na nasledujúcom obrázku je vidieť priamy kontakt s nečistotou s uvažovanou rezistivitou  $1 \cdot 10^{-4} \Omega \cdot m$ , ktorá predstavuje stav po dlhodobom používaní. Samozrejme, je veľmi obťažné deklarovať, že po akej dlhej dobe používania sa do takejto miery môže znečistiť povrch kontaktov. Pretože pri žiadnom predmete vyrobenom z rovnakého materiálu nedochádza k úplne identickému znečisteniu, degradácii alebo poškodeniu.



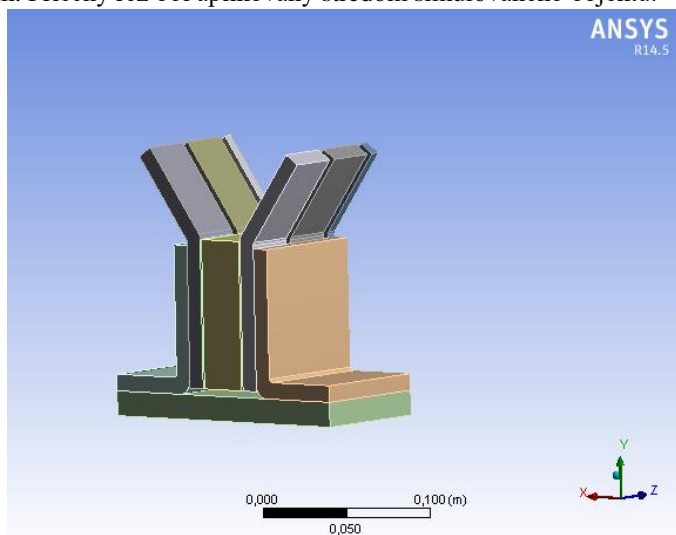
Obr. 4 Priebeh rozloženia teplôt priameho medeného kontaktu, s rezistivitou nečistoty  $1 \cdot 10^{-4} \Omega \cdot m$

Maximálna hodnota teploty na aktuálnom obrázku dosiahla  $118,82^\circ\text{C}$ , pričom minimálna teplota má hodnotu  $117,27^\circ\text{C}$ . Porovnaním týchto teplôt, s teplotami zistenými pri predchádzajúcich dvoch stavoch, sa preukázalo viac ako 2-násobné zvýšenie hodnôt. Oproti predošlému, „reálnemu“ prevádzkovému stavu, došlo k nárastu najvyššej teploty o  $73,476^\circ\text{C}$  a v porovnaní s „ideálnym“ prevádzkovým stavom sa najvyššia teplota zvýšila o  $79,778^\circ\text{C}$ .

Výrazné zmeny nastali aj z hľadiska rozloženia jednotlivých teplotných pásiem. Pri aktuálnom stave prevládajú pásma s vyššou teplotou (pásma s oranžovou a červenou farbou na obrázku), ktoré dokazujú, že najviac sa prehrievajú časti v blízkom okolí noža. Čiže znečistenie kontaktov môže vo veľkej miere ovplyvňovať teplotu elektrických spojov. Preto je dôležité vykonávať kontroly kontaktov v pravidelných intervaloch.

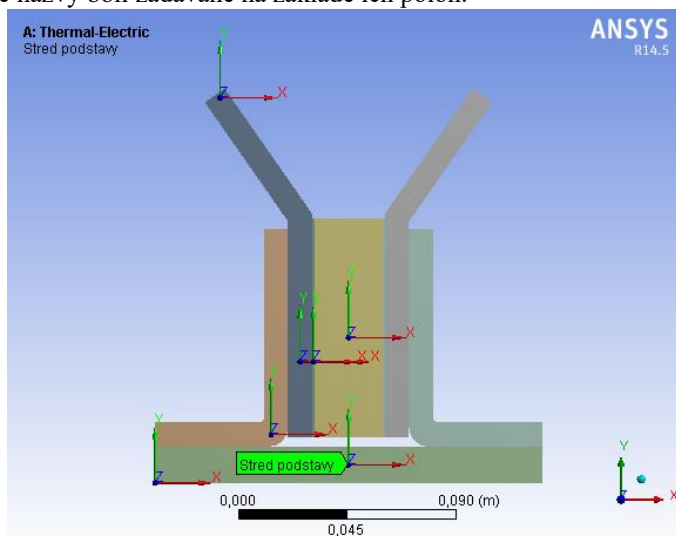
Ďalej sa uvažovalo, že najväčšie rozloženie maximálneho teplotného pásma sa nachádza v strede zostavy. Preto bola pre dôkladnejšiu analýzu vybraného kontaktu zvolený pohľad

s pričným rezom. Pričný rez bol aplikovaný stredom simulovaného objektu.



Obr. 5 Pričný rez vedený stredom priameho kontaktu

Po uskutočnení rezu boli na model umiestnené teplotné sondy („probes“). Tieto sondy umožňujú meranie teploty (diskrétneho elementu) v ktoromkoľvek bode. Celkovo bolo na rozličných miestach použitých 6 teplotných sond. Na Obr. 6 je znázornené ich rozmiestnenie, pričom sa nachádzajú na tých istých miestach ako súradnicové systémy. Podobne, ako je na obrázku uvedený názov teplotnej sondy v zelenej legende, tak sú pomenované aj ostatné teplotné sondy. Jednotlivé názvy boli zadávané na základe ich polôh.



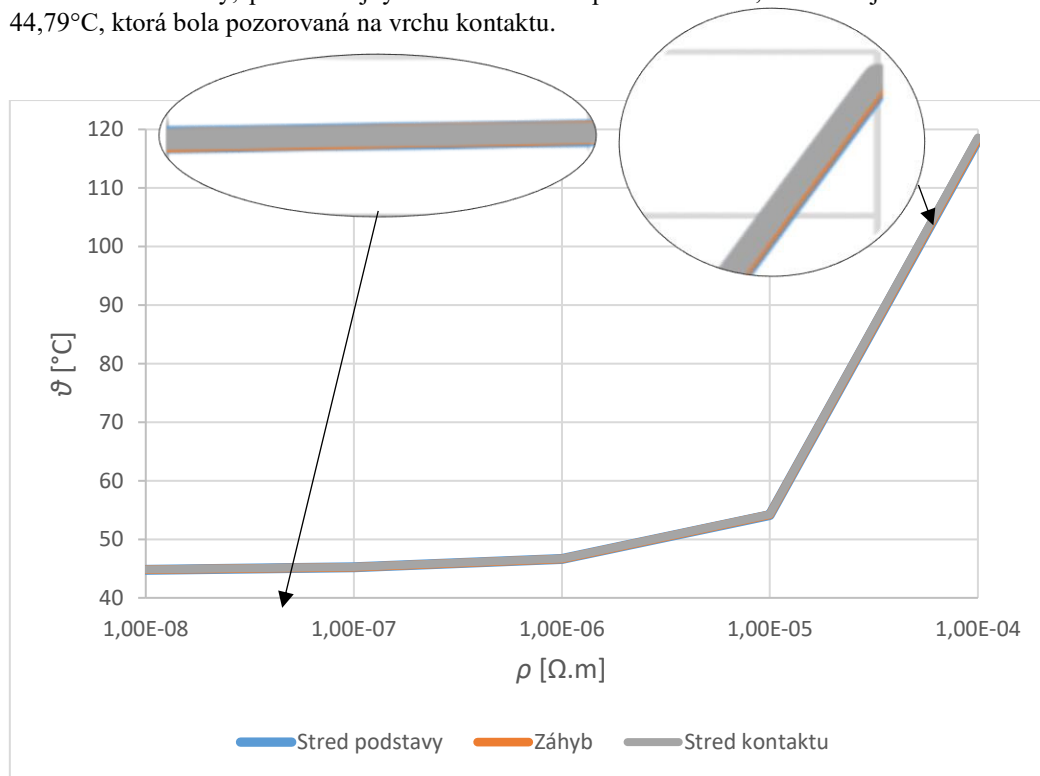
Obr. 6 Zobrazenie rozmiestnenia teplotných sond v strede priameho kontaktu

Hodnoty teplôt, ktoré boli namerané pomocou spomínaných teplotných sond, boli zapísané do tabuliek. Z nameraných hodnôt boli vytvorené grafické závislosti, ktoré poukazujú na to, ako sa mení teplota v určitých bodoch v strede kontaktu, pri pôsobení jednotlivých druhov znečistení.

Tabuľka 1  
 Namerané hodnoty teplotnými sondami, pri priamom kontakte z medi

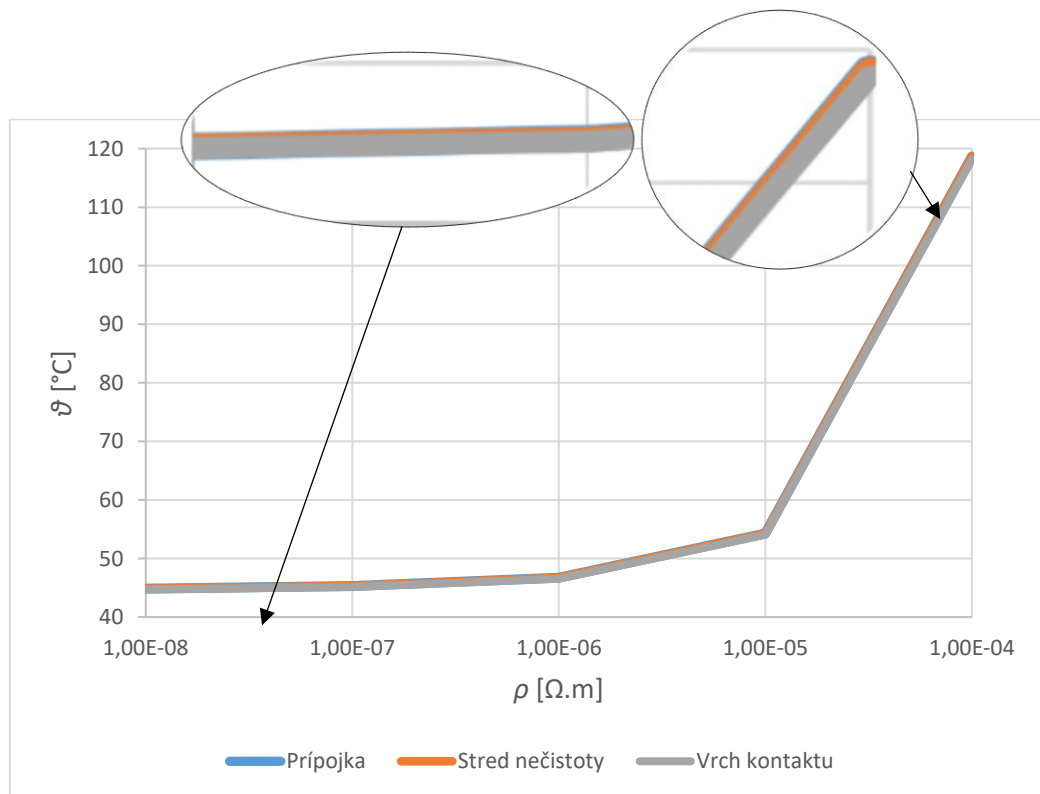
$\rho$ [ $\Omega \cdot m$ ]	Stred podstavy $\theta$ [ $^{\circ}C$ ]	Záhyb $\theta$ [ $^{\circ}C$ ]	Stred kontaktu $\theta$ [ $^{\circ}C$ ]	Stred noža $\theta$ [ $^{\circ}C$ ]	Stred nečistoty $\theta$ [ $^{\circ}C$ ]	Vrch kontaktu $\theta$ [ $^{\circ}C$ ]
$1 \cdot 10^{-4}$	118,21	118,37	118,62	118,66	118,67	118,11
$1 \cdot 10^{-5}$	54,133	54,175	54,236	54,246	54,243	54,066
$1 \cdot 10^{-6}$	46,632	46,659	46,696	46,7	46,698	46,564
$1 \cdot 10^{-7}$	45,215	45,239	45,272	45,276	45,274	45,147
$1 \cdot 10^{-8}$	44,79	44,813	44,845	44,849	44,847	44,723

Namerané teploty sondami, zobrazené v Tab. 1 nemajú výrazne rozdielne hodnoty, v niektorých prípadoch sa jedná maximálne o 0,5°C. Najvyššie teploty sa nachádzajú v strede noža alebo nečistoty, pričom najvyššia dosiahnutá teplota bola 118,67°C a najnižšia hodnota 44,79°C, ktorá bola pozorovaná na vrchu kontaktu.



Obr. 7 Závislosť teploty od rezistivity nečistôt pri priamom kontakte z medi

Kvôli lepšej prehľadnosti grafických závislostí boli zostrojené dva grafy, pričom do jedného aj druhého boli umiestnené 3 krivky. Poznámka: Pri uložení šiestich kriviek do jedného grafu by sa charakteristiky ešte viac prekryvali a nebol by viditeľný teplotný rozdiel.



Obr. 8 Závislosť teploty od rezistivity nečistôt pri priamom kontakte z medi

Z grafických závislostí na obrázkoch Obr. 7 a Obr. 8 je pozorovateľný nelineárny nárast teploty. Pri rezistivite nečistoty od  $1 \cdot 10^{-8} \Omega \cdot m$  do  $1 \cdot 10^{-6} \Omega \cdot m$  teplota narastala malým krokom po lineárnej krivke. Zmena začínala nastávať pri rezistivite nečistoty  $1 \cdot 10^{-6} \Omega \cdot m$  až  $1 \cdot 10^{-5} \Omega \cdot m$ , pričom najväčší skok teplôt bol medzi hodnotami  $1 \cdot 10^{-5} \Omega \cdot m$  až  $1 \cdot 10^{-4} \Omega \cdot m$ .

Táto skutočnosť nám potvrdzuje fakt, že medzi stredným a veľkým znečistením povrchu kontaktov je značná teplotná diferencia. Samozrejme, nárast teploty závisí aj od ďalších faktorov, medzi ktoré patrí ovplyvňovanie odvodu tepla z povrchu objektu. Teplotu môže ovplyvňovať napríklad ventilátor, ktorý ochladzuje elektrickú zostavu kontaktu (tzv. kabinet) alebo ventilátor, ktorý slúži na cirkuláciu ovzdušia v miestnosti a nachádza sa v blízkosti priameho kontaktu. Napriek tomu je dôležité jednotlivé kontakty v pravidelných intervaloch čistiť a ošetrovať.

#### IV. ZÁVER

Cieľom tohto článku bolo prezentovať výsledky rozloženia teploty na povrchu priameho kontaktu pri pôsobení nominálnej hodnoty prúdu 3000 A. V článku sú uvedené výstupy z programu ANSYS Workbench, pomocou ktorých bolo posudzované rozloženie teploty v okolí kontaktu. Simulácie preukázali nárast teplôt pri zvyšujúcich sa hodnotách rezistívít nečistôt. Tieto rezistivity predstavovali jednotlivé stupne znečistení, kde najväčším znečisteniam zodpovedali rezistivity s hodnotami  $1 \cdot 10^{-4} \Omega \cdot m$  a najmenším znečisteniam hodnoty rezistívít  $1 \cdot 10^{-8} \Omega \cdot m$ . Nárast teplôt potvrdili aj hodnoty namerané teplotnými sondami, ktoré boli rozmiestnené v rôznych častiach v strede priameho kontaktu. Grafické závislosti zostrojené z nameraných hodnôt následne vizualizovali nelineárny nárast teplôt pri zvyšujúcich sa rezistívítach nečistôt. Z uvedených výsledkov vyplýva, že je veľmi dôležité dbať na čistotu elektrických kontaktov, vykonávať pravidelné kontroly a ošetrovať povrch elektrických kontaktov.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576; ďalej Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu SR a SAV podporenej grantom VEGA 1/0435/19 a Kultúrnej a edukačnej grantovej agentúre MŠVVaŠ SR pod číslom KEGA 013TUKE-4/2021.

#### LITERATÚRA

- [1] KUČEROVÁ, E.: *Elektrotechnické materiály*. 1. vyd. Plzeň: Západočeská univerzita, 2002. 174 s. ISBN 80-7082-940-0.
- [2] DRÁPALA, J.; KURSA, M.: *Elektrotechnické materiály*. 1. vyd. Ostrava: VŠB – Technická univerzita Ostrava, 2012. 439 s. ISBN 978-80-248-2570-0.
- [3] PRESADA, J.: *Tepelná analýza usporiadania elektrických kontaktov*. Diplomová práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2021.



# Analýza výroby elektrickej energie pri zbernicovom systéme prehriatej pary

*Pavol SILAGYI, Dušan MEDVEĎ*

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

pavol.silagyi@student.tuke.sk, dusan.medved@tuke.sk

**Abstrakt** — Článok je venovaný výpočtovému nástroju, vytvorenému pre potreby spoločnosti U. S. Steel Košice, s.r.o. Zaoberá sa stručným popisom zariadení v podniku, súvisiacimi s výpočtovým nástrojom, popisom štruktúry výpočtového nástroja a samotným výpočtom žiadaných veličín. Predmetom práce bolo vytvorenie uceleného výpočtového nástroja, ktorého význam spočíva v správnom prerozdelení toku prehriatej pary v zbernicových systémoch na zariadenia podľa požiadaviek a následnom výpočte množstva elektrickej energie, ktorá je vyrábaná na turbogenerátoroch v podniku. Problém, spojený s vytváraním výpočtového nástroja, spočíva v komplexnosti prevádzky v podniku, kedy sa viacero zariadení recipročne ovplyvňuje. Účelom práce je spracovať údaje o jednotlivých zariadeniach tak, aby bolo možné vytvoriť prehľadný sumár pre prerozdelenie pary medzi jednotlivé zariadenia s ohľadom na požiadavky obsluhy a plánovania, ako aj prehľad o hodinových množstvách elektrickej energie vyrobenej na turbogenerátoroch. Výsledkom práce sú tak údaje o hodinovej výrobe elektrickej energie na turbogenerátoroch, maximálna a minimálna celková hodinová výroba elektrickej energie a prehľad rozdelenia pary na zariadenia.

**Kľúčové slová** — zbernica pary, zbernicový systém prehriatej pary, výpočtový nástroj, spoločná zberňa pary, výpočet elektrickej energie

## I. ÚVOD

Tepláreň vo firme U. S. Steel Košice patrí medzi najväčšie priemyselné teplárne, ktoré boli vybudované v bývalom Československu v rámci podniku. V dnešnej dobe ju tvorí päť kotlov, ktoré dokážu ako palivo spaľovať zemný plyn, uhlie a plyny, ktoré vznikajú v technologickom procese [1].

V rámci teplárne sú v časti „Strojovňa“ umiestnené turbogenerátory, ktoré slúžia na výrobu elektrickej energie, pričom sa značná časť pary, ktorá prechádza turbogenerátormi, využíva na pohon ďalších zariadení umiestnených v časti Strojovňa. Medzi zariadenia, ktoré sú poháňané parou z regulovaných odberov turbogenerátorov, patria napríklad turbodúchadlá, turbokompresory, vysokotlakové ohrievače a špičkové ohrievače slúžiace na výrobu tepla pre podnik a príslušné sídlisko Šaca [1].

Výroba elektrickej energie závisí, okrem chodu jednotlivých turbogenerátorov, taktiež aj od potreby pary pre zariadenia vyrábajúce médiá pre podnik. Preto je potrebné komplexne poznať prevádzku a mať prehľad o prevádzkovaných zariadeniach. Z týchto údajov je potom možné zistiť dosiahnuteľnú výrobu elektrickej energie pre potreby operátorov a dispečerov.

Predmetom tohto článku bolo zostaviť výpočtový nástroj v prostredí programu Microsoft Excel, ktorý je schopný vypočítať maximálnu a minimálnu možnú výrobu elektrickej energie pri definovanej dostupnosti zdrojov a pri konkrétnom rozdelení pary na zariadenia, ktoré sú parou poháňané. Tento nástroj bude následne používaný v praxi pre plánovanie prevádzky turbogenerátorov a predovšetkým plánovanie výroby elektrickej energie, typicky pre nasledujúci deň. Pri zostavovaní výpočtového nástroja boli využité poznatky, vychádzajúce zo skúseností operátorov, dispečerov a plánovačov.

## II. VÝPOČTOVÝ NÁSTROJ

Informácia o množstve elektrickej energie vyrobenej na turbogenerátoroch je dôležitá pre udržiavanie požadovanej rovnováhy medzi elektrickou energiou, ktorá je vyrobená a elektrickou energiou, ktorú podnik nakupuje, ako aj pre potreby poskytovania podporných služieb [2].

Množstvo elektrickej energie, vyrobenej na turbogenerátoroch v podniku závisí od potrebného

množstva pary na jednotlivých tlakových úrovniach. Zbernice pary 1,76 MPa, 0,63 MPa, 0,25 MPa ovplyvňujú turbogenerátory množstvom odoberanej pary, ktorá je odoberaná z regulovaných odberov turbogenerátorov, pričom tieto odbery z turbogenerátorov predstavujú primárne zdroje pary pre dané zbernicové systémy. Celková vyrobená elektrická energia na kondenzačno-odberových turbogenerátoroch sa vyrába kombináciou vyrobenej mechanickej energie z kondenzácie a z regulovaných odberov. Mierna zmena nastáva pri turbogenerátoroch TG3 a TG4, ktoré sú vyhotovené len ako protitlakové bez regulovaných odberov a ich celková výroba elektrickej energie závisí iba od množstva výstupnej pary [2], [3].

#### A. Požiadavky na výpočtový nástroj

Pri plánovaní nákupu a výroby elektrickej energie je žiadúce poznať aktuálne, avšak aj predpokladané, množstvo elektrickej energie, vyrobenej na turbogenerátoroch. Aktuálne hodnoty sú kontinuálne merané a údaje sa načítavajú na prehľadovú plochu, ktorá je dostupná v aktuálnom čase v rámci internej podnikovej siete, spolu s ostatnými údajmi pre ďalšie prevádzky – takzvaný Blackview [2].

Na výpočet predpokladaných hodnôt bolo však nutné zostrojiť výpočtový nástroj, ktorý by bral do úvahy všetky predvídateľné vplyvy, ktoré môžu ovplyvniť množstvo vyrobenej elektrickej energie. Najväčší vplyv na množstvo vyrobenej elektrickej energie majú zariadenia v podniku, presnejšie ich chod alebo odstávka, a množstvo pary, ktoré odoberajú. Každý výpadok týchto zariadení znamená menšie množstvo odoberanej pary, s čím súvisí aj množstvo vyrobenej elektrickej energie. Keď sa skumulujú potreby pary pre účely podniku, spolu so spotrebou pary pre zahĺtenie kondenzačných stupňov, vznikne celková potreba pary, ktorú musia parné kotly, ako zdrojové zariadenia, dodať. Výroba elektrickej energie je teda daná možnosťami odberu technologickej pary a aj možnosťami dodávky pary z parných kotlov [2].

Ďalšie vplyvy na výrobu elektrickej energie [2]:

- množstvo a zloženie paliva do kotlov a jeho výhrevnosť,
- obmedzenia množstva palív do horákov,
- technická kapacita odberov turbogenerátorov,
- množstvo potrebnej odoberanej pary (napr. vplyv ročného obdobia – vykurovanie),
- potrebné množstvá vyrábaných médií – stlačený vzduch a fúkaný vietor pre vysoké pece, para do podniku.

Okrem toho, treba brať do úvahy súvislosť jednotlivých zariadení so zariadeniami z celej prevádzky. Napríklad pri plánovanej odstávke vysokých pecí je nutné brať do úvahy odstavenie príslušného turbodúchadla. Taktiež množstvo potrebnej pary súvisí s ročným obdobím. Výsledkom nutnosti ohrevu vykurovacej vody v zimnom období je vyššia spotreba pary a teda je nutné na kotloch vyrobiť viac pary, ktorá následne prechádza lopatkami turbogenerátorov a vyrába elektrickú energiu [2], [3].

#### B. Štruktúra výpočtového nástroja

Pri tvorbe výpočtového nástroja bolo potrebné brať do úvahy všetky vplyvy, uvedené v predchádzajúcej časti článku. Základnou myšlienkou pri tvorbe programu bolo vychádzať z bilancovaných systémov pary 9,4 MPa, 1,76 MPa a 0,63 MPa. Boli vypracované v tabuľkovej forme prehľady o stavoch zariadení, meraných veličinách a ich obmedzeniach prípadne iných medziach. Celý výpočtový nástroj bol vytvorený v rozhraní systému Microsoft Excel [2].

Vzhľadom k myšlienke vychádzať z bilancovaných systémov sa ako s prvými začalo s kotlami a teda prvou tabuľkou je tabuľka pre kotly. Okrem aktuálneho stavu parných kotlov bolo nutné brať do úvahy taktiež druh, množstvo a výhrevnosť paliva vstupujúceho do kotlov. Výstupmi tejto tabuľky sú informácie o množstvách prehriatej pary vyrobenej jednotlivými kotlami pri danom palivovom zložení. Tieto výstupy sú napájaním spoločnej zbernice 9,4 MPa pary [2], [4], [5], [6], [7], [8].

Ďalšou z tabuliek je tabuľka, popisujúca stav jednotlivých turbogenerátorov spolu s obmedzeniami množstiev pary na jednotlivé odbery s upozornením na prekročenie určených medzí, v ktorých sa tieto množstvá pary môžu pohybovať. Tieto odbery sú zdrojmi pary pre jednotlivé zbernice o tlakových úrovniach 1,76 MPa, 0,63 MPa a 0,25 MPa. Výstupmi tabuliek sú množstvá vyrobenej elektrickej energie pre jednotlivé turbogenerátory [2], [9], [10], [11], [12], [13], [14].

Rovnako sa postupovalo aj pri ostatných zariadeniach, pričom pre lepšiu prehľadnosť výpočtového nástroja boli tabuľky s informáciami o zariadeniach rozdelené pomyselne na [2]: zdroje prehriatej pary pre jednotlivé zbernicové systémy; spotrebiče pary, odkiaľ para postupuje už iba do kondenzátora.

Údaje, ktoré sú v tabuľkách uvedené, okrem určených medzí, sú zadávané ručne alebo

načítavané z takzvaného DKEN. DKEN je Dátový koncentrátor energetiky, a ako taký, združuje signály o technologických parametroch energetických zariadení. Údaje sa do DKEN načítavajú pri každej zmene meranej veličiny a teda sa jedná o okamžité a presné hodnoty. V našom prípade sa však načítané údaje nemenia sekundovo automaticky, ale načítavajú sa jednorazovo na pokyn operátora pre potreby výpočtového nástroja. Údaje z DKEN slúžia na rýchle zadanie potrieb pary pre množstvo spotrebičov, podľa aktuálneho stavu alebo prípadne pre simulovaný výpočet pri simulovanom vypnutí niektorého zo zariadení vo výpočtovom nástroji, kedy operátor, ktorý bude vykonávať výpočet, nie je nútený zadávať hodnoty ručne, ale môže načítať aktuálne hodnoty a vypočítať prípadné množstvo vyrobenej elektrickej energie pri vypnutom alebo zapnutom zariadení alebo viacerých zariadeniach – nezávisle od skutočnosti [1], [2].

### III. DEMONŠTRÁCIA VYUŽITIA VÝPOČTOVÉHO NÁSTROJA NA SKUTOČNEJ UDALOSTI

Pre ďalšiu demonštráciu využitia výpočtového programu je popísaná udalosť, ku ktorej došlo dňa 2.12.2020. V ten deň boli naplánované odstávky na vysokej peci VP1 a zároveň odstávka spalínovej kotolne na oceliarni – SKOC, pričom vysoká pec VP3 bola odstavená už pred danou udalosťou. Spalínová kotolňa na oceliarni slúži na chladenie traktu odsávania spalín nad konvertormi, pričom dôjde k produkcii pary 1,76 MPa pre podnik. Okrem týchto odstávok bol mimo prevádzky aj parný kotol PK2. Z danej situácie vyplýva potreba zvýšenej výroby 1,76 MPa pary a zároveň výrazné ubratie vysokopecného plynu ako paliva na vstupe do kotlov [2].

Okrem týchto obmedzení bola v decembri 2020 nežiaduca prípadná odstávka turbogenerátora TG3V, ktorý je jedným z odberateľov pary 1,76 MPa. Táto požiadavka súvisí s ročnou dobou a s možným poškodením častí, súvisiacich s turbogenerátorom TG3V, ktoré sú uložené v priestoroch, vystavených vonkajším vplyvom, predovšetkým mrazu [2], [3]. Vzhľadom na požiadavky, boli vypracované tri alternatívy, ktoré boli následne prezentované zainteresovaným subjektom. Zároveň bola v tom istom čase prezentovaná aj najlepšia alternatíva [2].

#### A. Alternatíva č.1

V rámci alternatívy č. 1 sa počíta so zvýšením spotreby zemného plynu až na úrovni 12 tis. m<sup>3</sup>/h. Potreba zvýšenia spotreby pary súvisí s obmedzením vysokopecného plynu vzhľadom na chod iba jednej vysokej pece. Množstvo vysokopecného plynu pre tepláreň teda činí iba 60 tis. m<sup>3</sup>/h. K dispozícii pre tepláreň bolo v danom dni takisto aj 24 tis. m<sup>3</sup>/h koksárenského plynu, 18 tis. m<sup>3</sup>/h konvertorového plynu a vzhľadom na chod všetkých uhoľných mlynov, aj 25 t/h uhlia na kotloch PK6 a PK7. Plánované spotreby palív, spolu s predpokladanou výrobou pary na parných kotloch, je znázornená v tabuľke Tab 1 Spotreba palív a množstvo prehriatej pary pre alternatívu č. 1 [2].

Tabuľka 1  
 Spotreba palív a množstvo prehriatej pary pre alternatívu č. 1 [5], [7], [8], [2]

Bilancia pary	stav	Vyp. podľa palív	VPP	Koksárenský plyn	Konvertorový plyn	Zemný plyn	Energetické uhlie	Ent. 9,4 pary: 3,483	
		Vyrobená para						Napájacia voda	
		t/h	tis. m <sup>3</sup>	tis. m <sup>3</sup>	tis. m <sup>3</sup>	tis. m <sup>3</sup>	t/h	Teplota NV	Entalpia
Kotol K1	1	148	35	12,0	8,0	6,0		29	0,135
Kotol K2	0	0	0	0,0	0,0	0,0		197	0,845
Kotol K3	1	182	25	12,0	10,0	6,0		196	0,838
Kotol K6	1	231		0,0		0,0	25	204	0,877
Kotol K7	1	232		0,0		0,0	25	207	0,886
<b>Spolu</b>		<b>793</b>	<b>60</b>	<b>24</b>	<b>18</b>	<b>12</b>	<b>50</b>	<b>833</b>	<b>3,580</b>

Chybové hlášky:

Pri tejto alternatíve nebol obmedzený chod ani jedného zo zariadení, čo znamená, že okrem samozrejme vyradenia turbodúchadiel TD1 a TD3, ktoré súvisia s výrobou fúkaného vzduchu do vysokých pecí VP1 a VP3, toho času odstavených, nebolo z prevádzky vyradené ani jedno zariadenie. Naopak, bolo nutné zvýšiť zaťaženie odberu do podniku z vývodov 1 až 3 [2].

Maximálne a minimálne množstvo elektrickej energie, ktoré je možné pri danej alternatíve vyrobiť, je uvedené v tabuľke Tab. 2 Minimálne a maximálne množstvo elektrickej energie vyrobenej pre alternatívu č. 1 [2].

Tabuľka 2  
 Minimálne a maximálne množstvo elektrickej energie vyrobenej pre alternatívu č. 1 [2]

Maximálny dosiahnuteľný výkon:	
<b>100,97</b>	<b>MW</b>
Minimálny dosiahnuteľný výkon:	
<b>88,89</b>	<b>MW</b>

B. Alternatíva č.2

V alternatíve č. 2 sú množstvá vstupných palív na rovnakej úrovni. K rozdielu v množstvách palív dochádza v prípade zemného plynu, ktorý bude znížený z pôvodných 12 tis. m<sup>3</sup>/h na 9 tis. m<sup>3</sup>/h. Zníženie množstva zemného plynu je dôležité pri predchádzaní prekročenia denného obmedzenia spotreby zemného plynu v podniku. Množstvá palív na vstupe do parných kotlov a množstvo pary, ktoré je možné pri danom množstve paliva dosiahnuť, je uvedené v tabuľke Tab. 3 Spotreba palív a množstvo prehriatej pary pre alternatívu č. 2 [2].

Tabuľka 3  
 Spotreba palív a množstvo prehriatej pary pre alternatívu č. 2 [4], [5], [6], [7], [8], [2]

Bilancia pary	Stav	Vyp. podľa palív	VPP	Koksárenský plyn	Konvertorový plyn	Zemný plyn	Energetické uhlie	Ent. 9,4 pary: 3,483	
		Vyrobená para						Napájacia voda	
		t/h	tis. m <sup>3</sup>	tis. m <sup>3</sup>	tis. m <sup>3</sup>	tis. m <sup>3</sup>	t/h	Teplota NV	Entalpia
Kotel K1	1	139	35	12,0	8,0	5,0		29	0,135
Kotel K2	0	0	0	0,0	0,0	0,0		197	0,845
Kotel K3	1	160	25	12,0	10,0	4,0		196	0,838
Kotel K6	1	231		0,0		0,0	25	204	0,877
Kotel K7	1	232		0,0		0,0	25	207	0,886
<b>Spolu</b>		<b>762</b>	<b>60</b>	<b>24</b>	<b>18</b>	<b>9</b>	<b>50</b>	<b>833</b>	<b>3,580</b>

Chybové hlásky:

Oproti alternatíve č. 1, kde sa nepočítalo s odstávkou zariadení, okrem tých nevyhnutných, v prípade alternatívy č. 2 sa počíta s odstávkou turbokompresora TK1. Stratu množstva stlačeného vzduchu, ktorý je vyrábaný na turbokompresore TK1 nahradíme zapojením elektrických kompresorov EK1 a EK2 do prevádzky. Týmto zabezpečíme zníženie odoberaného množstva 1,76 MPa pary spotrebičmi 1,76 MPa pary o 30 t/h [2].

Výroba elektrickej energie na kondenzačných turbogenerátoroch je teda približne rovnaká ako v prípade alternatívy č. 1. Rozdiel nastáva v prípade turbogenerátorov TG3 a TG4 a to z dôvodu obmedzenia odberu 1,76 MPa. Maximálne a minimálne množstvo elektrickej energie, ktorú je možné pri danej alternatíve vyrobiť, je vyjadrená v tabuľke Tab. 4 Minimálne a maximálne množstvo elektrickej energie vyrobenej pre alternatívu č. 2 [2].

Tabuľka 4  
 Minimálne a maximálne množstvo elektrickej energie vyrobenej pre alternatívu č. 2. [2]

Maximálny dosiahnuteľný výkon:
<b>98,28 MW</b>
Minimálny dosiahnuteľný výkon:
<b>85,78 MW</b>

C. Alternatíva č.3

Alternatíva č. 3 spočíva v ďalšom znižovaní množstva zemného plynu na vstupe do parných kotlov. Množstvo zemného plynu na vstupe do kotlov je znížené na 6 tis. m<sup>3</sup>/h. Množstvo zvyšných palív na vstupe do kotlov ostáva na pôvodných hodnotách tak, ako tomu bolo aj v predošlých alternatívach. Množstvá palív pre alternatívu č. 3 sú vyjadrené v tabuľke Tab. 5 Spotreba palív a množstvo prehriatej pary pre alternatívu č. 3 [2].

Tabuľka 5  
 Spotreba palív a množstvo prehriatej pary pre alternatívu č. 3 [4], [5], [6], [7], [8], [2]

Bilancia pary	Stav	Vyp. podľa palív	VPP	Koksárenský plyn	Konvertorový plyn	Zemný plyn	Energetické uhlie	Ent. 9,4 pary: 3,483	
		Vyrobená para						Napájacia voda	
		t/h	tis. m <sup>3</sup>	tis. m <sup>3</sup>	tis. m <sup>3</sup>	tis. m <sup>3</sup>	t/h	Teplota NV	Entalpia
Kotel K1	1	121	35	12,0	8,0	3,0		29	0,135
Kotel K2	0	0	0	0,0	0,0	0,0		197	0,845
Kotel K3	1	148	25	12,0	10,0	3,0		196	0,838
Kotel K6	1	231		0,0		0,0	25	204	0,877
Kotel K7	1	232		0,0		0,0	25	207	0,886
<b>Spolu</b>		<b>733</b>	<b>60</b>	<b>24</b>	<b>18</b>	<b>6</b>	<b>50</b>	<b>833</b>	<b>3,580</b>

Chybové hlásky:

Alternatíva č. 3 je oproti alternatívam č. 2 a č. 1 „agresívnejšia“ nielen z pohľadu obmedzenia zemného plynu, ale aj z hľadiska obmedzenia chodu strojov. Okrem odstávky turbokompresora TK1 bude odstavený aj turbogenerátor TG3V. Vzhľadom na to, že turbogenerátor TG3V je

jedným zo spotrebičov pary 1,76 MPa, jedná sa o úsporu pary pre zabezpečenie dostatočného množstva pary pre parný systém 1,76 MPa [2].

Množstvo pary, ktoré je dodávané od parných kotlov do systému, je zo všetkých alternatív najmenšie. Aj napriek zníženiu spotreby pary je nutné ďalšie ubratie pary na kondenzáciách a to v množstve 66,91 t/h. Spolu so zníženým prietokom pary v systéme to spôsobí pokles výroby elektrickej energie. A to nie iba v prípade kondenzačných turbogenerátorov TG1, TG2 a TG5, ale tak isto aj v prípade pretlakových turbogenerátorov TG3 a TG4. Okrem toho, je odstavený aj turbogenerátor TG3V, čo spôsobí celkový pokles výroby elektrickej energie. Celkové množstvo elektrickej energie, ktorú je možné vyrobiť, je vyjadrené v tabuľke Tab. 6 Minimálne a maximálne množstvo elektrickej energie vyrobenej pre alternatívu č. 3 [2].

Tabuľka 6  
Minimálne a maximálne množstvo elektrickej energie vyrobenej pre alternatívu č. 3 [2]

Maximálny dosiahnuteľný výkon:	
<b>91,59</b>	<b>MW</b>
Minimálny dosiahnuteľný výkon:	
<b>80,49</b>	<b>MW</b>

#### D. Výber vhodnej alternatívy

Posledným krokom pri plánovaní je výber najvhodnejšej alternatívy. Táto alternatíva by mala spĺňať čo najviac požiadaviek a to [2]:

- znížený odber zemného plynu,
- zachovanie chodu turbogenerátora TG3V,
- ohľad na chod iba jedinej vysokej pece a s tým spojené množstvo vysokopecného plynu na vstupe do kotlov,
- ohľad na zvyšné dostupné množstvá plynov na vstupe do kotlov.

Boli navrhnuté tri alternatívy, ktoré sú vyjadrené v tabuľke Tab. 7 Porovnanie alternatív č. 1, 2 a 3 pre výber najvhodnejšej alternatívy [2].

Tabuľka 7  
Porovnanie alternatív č. 1, 2 a 3 pre výber najvhodnejšej alternatívy [2]

Alternatíva	Akcia	Výroba pary	Spotreba ZP	Výroba EE
1	Bez odstávky zariadení, zvýšenie množstva ZP	793 t/h	12 tis.m3/h	100,97 MW
2	Odstávka TK1, nábeh EK1 a EK2	<b>762 t/h</b>	<b>9 tis.m3/h</b>	<b>98,28 MW</b>
3	Odstávka TG3V, odstávka TK1, nábeh EK1 a EK2	733 t/h	6 tis.m3/h	91,59 MW

Aj napriek tomu, že v prípade použitia alternatívy č. 1 je výroba elektrickej energie najvyššia, najvyššia je tak isto aj spotreba zemného plynu. Pri spotrebe zemného plynu je potrebné byť opatrný, pretože môže dôjsť k prekročeniu dennej medze spotreby zemného plynu, čo by viedlo k značným pokutám, ktorým sa spoločnosť snaží vyhnúť. K prekročeniu nemusí dôjsť výhradne spaľovaním plynu v plynových parných kotloch, ale tak isto je nutné brať ohľad na spaľovanie v iných častiach podniku, kedy by súčet spotrieb mohol prekročiť dennú medzu spotreby a nastal by nežiaduci stav. Preto je nutné, v prípade zemného plynu, vhodne narábať s jeho spotrebou. Alternatíva č. 1 je z toho dôvodu nevhodná na dlhodobé použitie, s ktorou sa v prípade plánovania uvažuje [2].

Alternatíva č. 2 je z pohľadu požiadaviek, a tak isto aj z pohľadu spotreby zemného plynu, najlepším kompromisom zo všetkých vyššie spomenutých. Aj napriek nábehu elektrických kompresorov EK1 a EK2, ktoré nahrádzujú odstavený turbokompresor TK1, a s tým spojená vyššia spotreba elektrickej energie v podniku, čo v prípade EK1 a EK2 predstavuje 6 MW, je v značnej časti prípadov výhodnejšie nakúpiť potrebnú elektrickú energiu ako spotrebovať zemný plyn. Ak k tomu pripočítame možnosť pokuty za prekročenie spotreby zemného plynu pri nadmernej spotrebe, je nákup elektrickej energie lacnejšou alternatívou [2].

Alternatíva č. 3 je prispôbená možnosti, kedy by bola možná iba minimálna spotreba zemného plynu, pričom by bol odstavený turbogenerátor TG3V za účelom zabezpečenia

dostatočného množstva pary pre zbernicový systém 1,76 MPa. Táto alternatíva je ale nevhodná z hľadiska krátkodobého odstavenia turbogenerátora TG3V a s tým spojených poškodení častí vystavených mrazom [2].

Pre daný deň bola vybratá alternatíva č. 2. Táto alternatíva splnila očakávania a po jej zrealizovaní sa ukázalo, že nepresnosť výpočtu výroby elektrickej energie bola nižšia ako 2 MW, čo vo vzťahu k plánovaniu predstavuje zanedbateľný rozdiel a považuje sa to za presný výpočet. Zároveň bola zachovaná dostatočná rezerva zemného plynu pri danej spotrebe na parných kotloch [2].

#### IV. ZÁVER

Z tohto článku vyplýva, že prevádzka teplárne v podniku U. S Steel Košice sa líši od prevádzok v klasických teplárnach, čo v konečnom dôsledku sťažuje výpočet výroby elektrickej energie. Výroba elektrickej energie na jednotlivých turbogenerátoroch je závislá nielen od výroby prehriatej pary, ktorá je dodávaná do systému, ale najmä od jednotlivých odberov z turbogenerátorov. Tieto odbery sú ovplyvnené chodom zariadení v prevádzke teplárne, ktorých chod je závislý od chodu zariadení, ktoré nemusia byť súčasťou teplárne. Ako príklad nám slúži prevádzka vysokých pecí a jej vplyv na množstvo palív na vstupe do parných kotlov a zároveň vplyv na chod turbodúchadiel v prevádzke teplárne. A preto bolo nutné brať ohľad nielen na prevádzku teplárne, ale takisto aj na chod zariadení v iných prevádzkach.

Výpočtový program, popísaný v tomto článku, bol odskúšaný ako plánovačmi, tak obsluhou. Výsledkom je zaradenie tohto programu do bežnej praxe pri plánovaní prevádzky na nasledujúce obdobie. Je preto možné povedať, že výpočtový program splnil očakávania a vstupné požiadavky. Ďalším rozšírením programu o ďalšie funkcionality by bola zvýšená flexibilita a využiteľnosť výpočtového programu vo vzťahu k podniku ako celku a nie iba vo vzťahu k prevádzke teplárne.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu SR a SAV podporenej grantom VEGA 1/0757/21.

#### LITERATÚRA

- [1] Intranet spoločnosti U. S. Steel Košice. (cit. 2021-5-9). Dostupné na < [www.intranet.sk.uss.com](http://www.intranet.sk.uss.com) > [interná sieť spoločnosti U. S. Steel Košice s.r.o].
- [2] Belej, J. Vedúci oddelenia Energetické bilancie a stratégia. Konzultácie v priebehu časového obdobia od 2.2020 do 3.2021. (cit. 2021-5-9)
- [3] Pracovníci energetického dispečingu. Konzultácie v priebehu časového obdobia od 2.2020 do 10.2020. (cit. 2021-5-9)
- [4] Sadvár, D. Pracovná smernica II. úrovne pre prevádzku, obsluhu a udržiavanie parného kotla PK1 (DZ Energetika / Prevádzka Kotolňa). 2009. (cit. 2021-5-9). PS/DW10/07/09
- [5] Pristáš, J. PS pre kotol 2 Pre prevádzku, obsluhu a údržbu. 4.2.2011. (cit. 2021-5-9). PSIIZ/DW/0054
- [6] Sadvár, D. Pracovná smernica II. úrovne pre prevádzku, obsluhu a údržbu parného kotla PK3 (DZ Energetika / Prevádzka Kotolňa). 2009. (cit. 2021-5-9). PSIIZ/DW/0053
- [7] Pristáš, J. Pracovná smernica II. úrovne pre VTZ - Pracovná smernica kotla PK7 Pre prevádzku, obsluhu a údržbu. 20.6.2017. (cit. 2021-5-9). PSIIZ/FE/0001
- [8] Pristáš, J. Pracovná smernica II. úrovne pre VTZ - Pracovná smernica kotla PK7 Pre prevádzku, obsluhu a údržbu. 20.6.2017. (cit. 2021-5-9). PSIIZ/FE/0001
- [9] Hegedúš, E. Pracovná smernica II. úrovne - Turbogenerátor č.1 Platí pre obsluhu TG. 25.5.2016. (cit. 2021-5-9). PSII/DW/0009
- [10] Hegedúš, E. Pracovná smernica II. úrovne - Turbogenerátor č.2 Platí pre obsluhu TG. 19.6.2015. (cit. 2021-5-9). PSII/DW/0012
- [11] Hegedúš, E. Pracovná smernica II. úrovne - Turbogenerátor č.3 Platí pre obsluhu TG. 25.5.2011. [cit. 2020-5-13]. PSII/DW/0014
- [12] Hegedúš, E. Pracovná smernica II. úrovne - Turbogenerátor č.4 Platí pre obsluhu TG. 1.12.2011. [cit. 2020-5-18]. PSII/DW/00152
- [13] Hegedúš, E. Pracovná smernica II. úrovne - Turbogenerátor č.5. 9.2.2015. [cit. 2020-5-18]. PSII/DW/0028
- [14] Hegedúš, E. Pracovná smernica II. úrovne - Turbogenerátor TG3V. 16.12.2013. [cit. 2020-5-18]. PSII/DW/0019
- [15] Silagyí, P.: Analýza výroby elektrickej energie pri zbernicovom systéme prehriatej pary. Diplomová práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2021.

# Impedance analysis of inhibited mineral oil

<sup>1</sup>Peter Havran, <sup>2</sup>Roman Cimbala, <sup>3</sup>Vladimír Kohan

<sup>1</sup> Department of Electric Power Engineering, 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>3</sup> Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical university of Košice, Slovak republic

<sup>1</sup>peter.havran@tuke.sk, <sup>2</sup>roman.cimbala@tuke.sk, <sup>3</sup>vladimir.kohan@tuke.sk

**Abstract** — The essence of this paper is the analysis of inhibited mineral oil through complex impedance at different applied temperatures. The theoretical part explains the behaviour of dielectrics and relaxation processes through a complex impedance. The experimental part describes the description of the experiment and points out the dependences of the complex impedance of mineral oil in the increasing temperature.

**Keywords** — complex impedance, complex plane, impedance analysis, mineral oil

## I. INTRODUCTION

When monitoring the events that take place in insulating materials, it is also possible to use measurements of electrophysical parameters, the data of which are closely related to the structure of the material. A promising and often underappreciated method is also impedance analysis for measuring electrical and dielectric properties, which with its simplicity, speed and accuracy can be compared with proven methods in materials science. The sensitivity of impedance analysis is relatively high and universal when examining a wide range of materials of varying complexity [1]. Mineral oils are a major component of the electrical insulation system in various types of electrical equipment. In addition to their insulating properties, they also have a cooling effect. The cooling effect is very important in terms of increasing the life and reliability of the equipment. The current time is facing the challenge of lowering the operating temperature of the device, so scientists are looking for additives or new materials that improve the thermally conductive properties while maintaining the best possible insulation properties. Based on this, the investigation of new materials must have an analysis of the dielectric response in a wide frequency range, through a complex impedance varying with temperature and an applied electric field to determine the nature of the loss mechanism [2]. The motivation of this paper is to describe the behaviour of inhibited mineral oil in an electric field when changing the applied temperature by impedance analysis.

## II. COMPLEX IMPEDANCE

Impedance is a complex quantity, which is given by the sum of its real and imaginary part:

$$Z^* = Z' + jZ'' \quad (1)$$

The behaviour of the dielectric in a time-varying electric field can be analysed through a Voigt element, which represents a parallel RC circuit (Fig. 1a). Its impedance is given by [6]:

$$Z^* = \frac{R \cdot (1 - j\omega \cdot \tau)}{1 + \omega^2 \cdot \tau^2} \quad (2)$$

when  $\tau = R \cdot C$  is a time of relaxation. In Fig. 1b is a graphical representation of the impedance of a Voigt element, the plotting of which in the complex plane characterizes a frequency-dependent semicircle with a center on  $Z'$ , where  $\omega \rightarrow \infty$  intersects the point 0 [3]. The ideal case of a Voigt element characterizes in this figure the relaxation time  $\tau_1$ . If the values of capacitance  $C$  in the circuit gradually increase while maintaining the original values of  $R$ , the relaxation time gradually increases to  $\tau_2$  and  $\tau_3$  and the maximum of the curves decreases with deformation of their left side representing the area of higher frequencies [4].

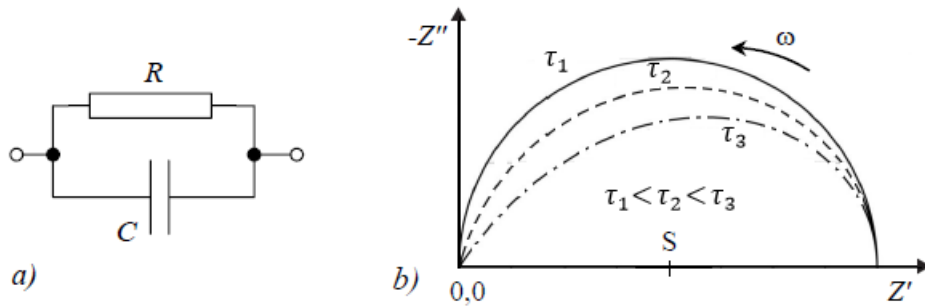


Fig. 1 The impedance of a Voigt element and its representation in a complex plane [4]

### III. EXPERIMENT

The measurement was performed in laboratory conditions, which included an IDAX 300 measuring instrument, a laptop with IDAX 4.1 software, a Tettex AG Zürich type 2965 temperature controller, a Tettex AG Zürich type 2903 heating device, thermometer, and a Tettex AG Zürich type 2903a (116 275) electrode system with a sample of inhibited mineral oil. The assembly of the measuring workplace is shown in Fig. 2.

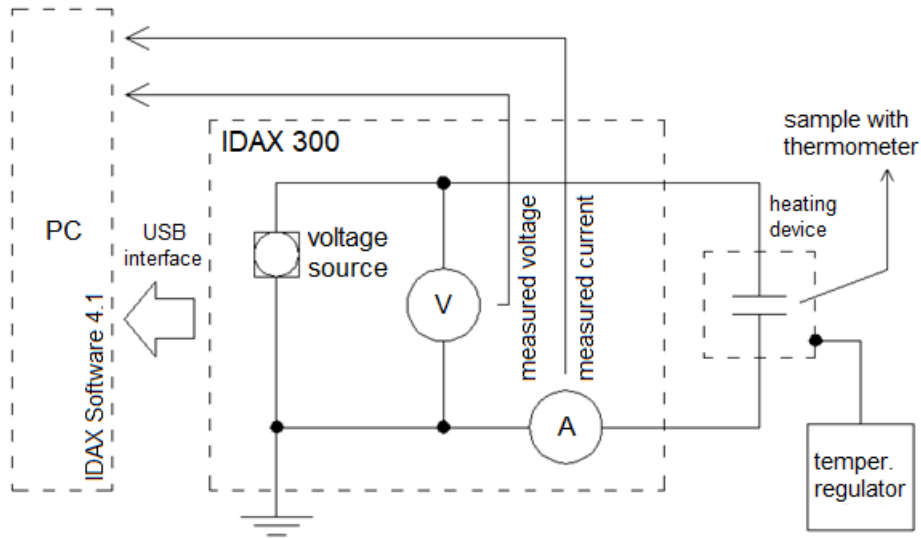


Fig. 2 Measuring workplace connection

The subject of the investigation was a sample of inhibited mineral oil MOGUL TRAF0 CZ-A (hereinafter "MOGUL") from the manufacturer Paramo. It is used as an insulating and cooling liquid of transformers of all voltage levels, machines with the highest voltage and power, usable also in switches, contactors, capacitors, and the like. Its characteristic properties include low density, high surface tension, excellent electrical insulating properties, high oxidative stability, and long service life [5].

MOGUL oil was measured at a gradually increasing temperature from 20 °C to 90 °C with a thermal step of 10 °C, which was adjusted using a temperature controller. All measurements were performed in the frequency band from 1 mHz to 3 kHz of the IDAX 300 measuring instrument at a set applied voltage of 100 V [7].

### IV. IMPEDANCE ANALYSIS

Using a measuring device, a voltage of 100 V was applied to a sample of MOGUL oil, where the system measured the current flowing through the material, from which the complex impedance was subsequently calculated, recorded in the IDAX 4.1 software in a portable laptop. Fig. 3 shows a graphical dependence of the real part of the complex impedance on the frequency at different thermal stresses, where it can be clearly seen that the values of the real part of the complex impedance decreased with increasing temperature in the low-frequency band. The area of low frequencies is known for the manifestation of conductivity processes taking place in the material. As the temperature increases, the viscosity decreases, and the conductivity in liquid materials increases. Around the frequency of 1 Hz, there was an opposite dispersion of the real part of the complex impedance, when its values increased with increasing temperature.



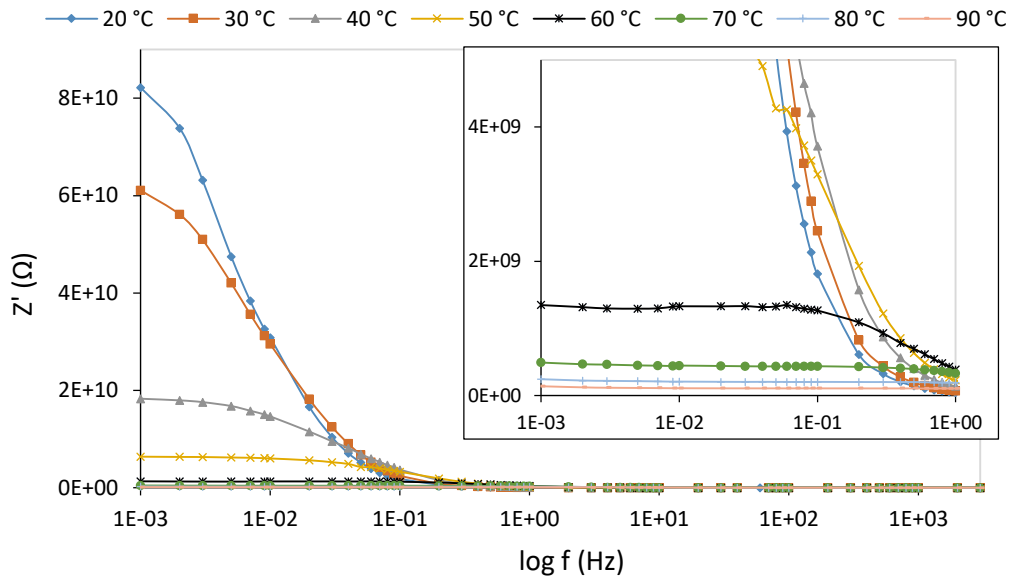


Fig. 3 Real part of the complex impedance in the measured frequency spectrum at different temperatures

The higher frequency range confirms the above-mentioned dispersion, where it is clear that the rapidly changing intensity of the applied electric field causes an increase in the real part of the complex impedance with increasing temperature, as shown in Fig. 4. From Fig. 3 and Fig. 4 it is possible to express the behaviour of the real part of the complex impedance through a decreasing trend with increasing frequency.

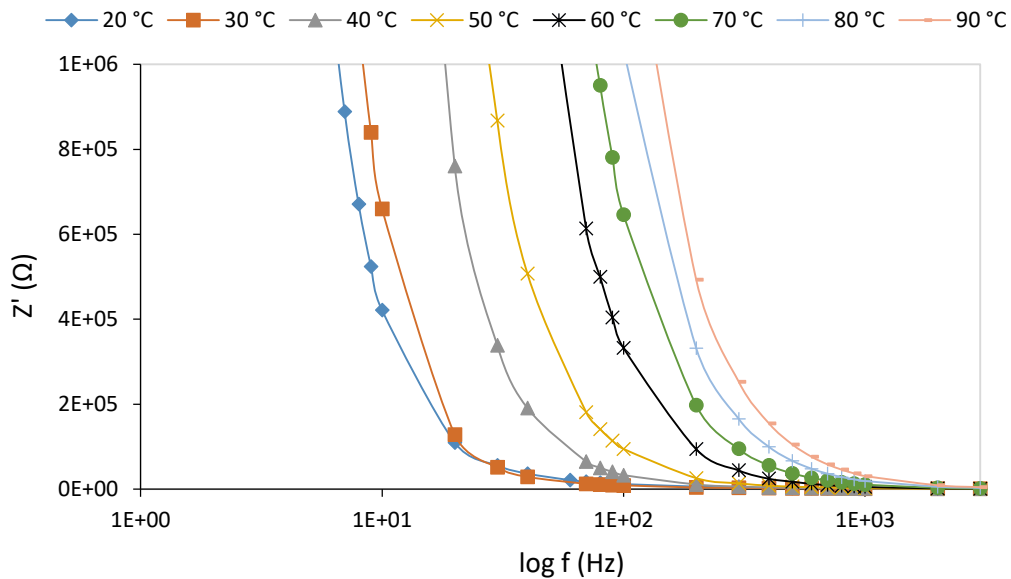


Fig. 4 Band of higher frequencies of the real part of the complex impedance at different temperatures

Fig. 5 shows the graphical dependence of the imaginary part of the complex impedance on the frequency at different thermal stresses. The imaginary part of the complex impedance characterizes dielectric losses in the material, so its course resembles the behaviour of the imaginary part of the complex permittivity with hills, which resemble polarization processes with certain relaxation times. From Fig. 5 it can be seen that the increasing oil temperature causes the peaks of the imaginary part of the complex impedance to decrease with their shift towards higher frequencies. Higher applied temperature causes deterioration of dielectric properties, resulting in a lower value of the imaginary part of the complex impedance in the frequency domain.

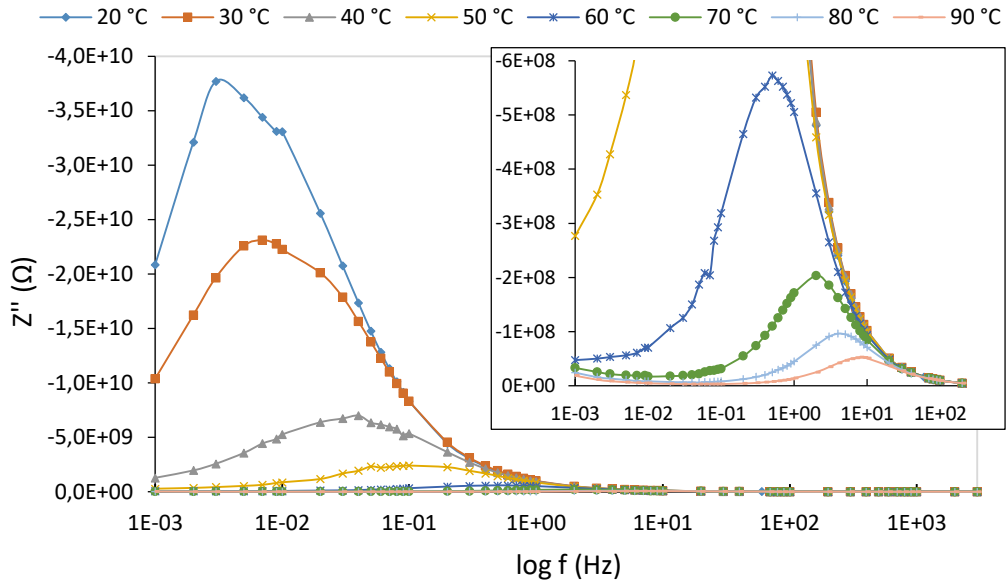


Fig. 5 Imaginary part of the complex impedance in the measured frequency spectrum at different temperatures

By plotting the impedance in the complex plane in Fig. 6 and Fig. 7, it is possible to see waveforms resembling the Cole-Cole distribution of relaxation times with waveforms in the shape of an imperfect semicircle. With increasing temperature, it is possible to see an improvement in the shape of the waveforms, which more closely resemble a perfect semicircle. This is related to the smaller distribution of relaxation times of the polarization mechanisms with increasing temperature, as shown in Fig. 1. Impedance values also decrease with a similar trend as in Fig. 3 and Fig. 5. The non-development of impedance to a quasi-circle at 20 °C and 30 °C is related to the insufficiently low measured frequency band, similarly to the dependence of the real and imaginary part of the complex impedance on frequency. Bringing the impedance to the complex plane confirmed that with increasing temperature the conductivity of mineral oil increases, when so-called "tails" in the low-frequency band were captured at 70 °C to 90 °C along with the semicircles, which are related to the increased conductivity of the material.

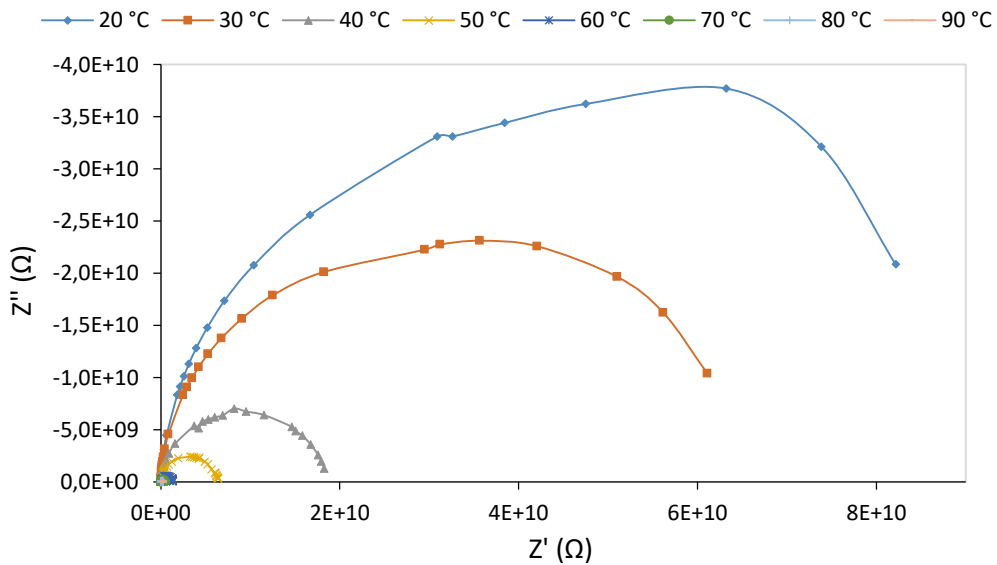


Fig. 6 Temperature dependence of MOGUL oil impedance in the complex plane

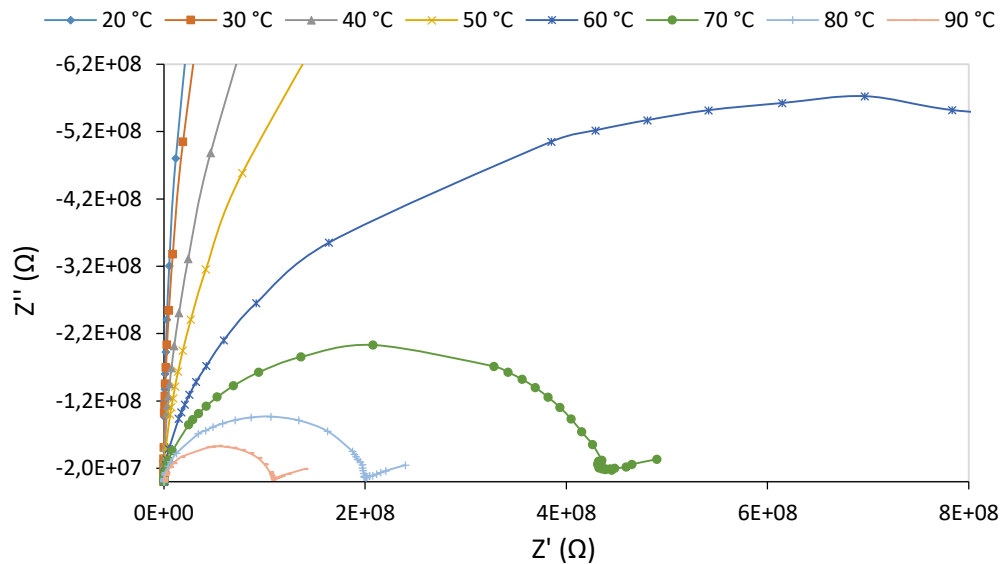


Fig. 7 Detail of the temperature dependence of the MOGUL oil impedance in the complex plane

## V. CONCLUSION

This paper aimed to investigate and analyse the inhibited mineral oil MOGUL in a wide frequency range, through a complex impedance varying with the temperature at the applied electric field to determine the nature of the loss mechanism. The real part of the complex impedance decreased in the low-frequency band up to 1 Hz with increasing temperature. From this frequency there was the opposite effect, i.e. in the area of higher frequencies, the real part of the complex impedance acquired higher values at a higher applied temperature. The increasing temperature of the oil causes the peaks of the imaginary part of the complex impedance to decrease with their shift towards higher frequencies. As the temperature increases, one can see an improvement in the shape of the waveforms, which more closely resembles a perfect Debye semicircle. This is related to the smaller distribution of relaxation times of polarization mechanisms with increasing temperature. Plotting the impedance to the complex plane confirmed that with increasing temperature the conductivity of mineral oil increases when so-called "tails" in the low-frequency band were captured at a temperature of 70 °C to 90 °C alongside the semicircles.

## VI. DISCUSSION

Impedance analysis of a dielectric material, which describes its behavior at the applied intensity of an electric field with a changing thermal field through a complex impedance, can be denoted by impedance spectroscopy. The dielectric has its impedance expression, so it is possible to describe events in insulating materials using complex impedance in the frequency spectrum, not only using commonly used dielectric parameters such as complex permittivity, complex capacitance or complex electrical modulus. The content of this work is a description of the behavior of a liquid dielectric in an electric field, which is stressed by temperatures from 20 °C to 90 °C, and pointing out its changes through a complex impedance, which has sufficient predictability in comparison with commonly monitored dielectric parameters in insulating materials in the degradation process.

## ACKNOWLEDGMENT

The work was supported by the Ministry of Education, Science, Research and Sport within the project VEGA 2/0011/20 and 1/0154/21 and the Slovak Agency for Research and Development based on contract no. APVV-15-0438, APVV-17-0372, and APVV-18-0160.

## REFERENCES

- [1] F. Vadim, „Impedance spectroscopy: Applications to Electrochemical and Dielectric Phenomena“, New Jersey: John Wiley & Sons, Inc. 2012, 356 pp. ISBN 978-0-470-62778-5.
- [2] M. Rajňák et al., „Temperature Dependence of a Dielectric Relaxation in Weakly Polar Ferrofluids“, Acta Physica Polonica A, Vol. 131, No. 4, 2017, pp. 943-945.

- [3] R. Cimbala, P. Semančík, "Tepelná degradácia kvapalných izolačných materiálov", Technical university of Košice, Letná 9, Košice. 2016. 147 pp. ISBN 978-80-553-2545.
- [4] R. Schmidt, „Ceramic Materials Research Trends: Impedance Spectroscopy of Electroceramics“, Nova Science Publishers, Inc. 2007, 321-351 pp. ISBN 978-1-60021-769-2.
- [5] PARAMO, MOGUL TRAF0 CZ-A, Technical Data Sheet. Online: <[https://eshop.paramo.cz/data/VyrobkovaDokumentace/ti-trafo\\_cza\\_z5.pdf](https://eshop.paramo.cz/data/VyrobkovaDokumentace/ti-trafo_cza_z5.pdf)>.
- [6] J. R. Macdonald, „Impedance spectroscopy“, In. Annals of Biomedical Engineering. 1991, 20, 2 pp. ISSN 1573-9686.
- [7] P. Havran, R. Cimbala, „Dielektrická spektroskopia online tepelne namáhaného kvapalného dielektrika“, Electrical Engineering and Informatics XI, Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, Letná 9, Košice. 2020. 173-178 pp. ISBN 978-80-553-3585-8.

# Kódovanie ľudskej hlavy v štandarde MPEG-4 SNHC

Ján MIHALÍK

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

jan.mihalik@tuke.sk

*Abstract* — SNHC (Synthetic Natural Hybrid Coding) je podskupinou MPEG-4 [4], ktorá sa špecializuje na kódovanie grafických modelov reálnych alebo virtuálnych 3R objektov. Štandardizácia tohoto kódovania rozširuje rozsah pôvodných aplikácií MPEG-4, pretože umožňuje kombinovať reálne a syntetické objekty vo virtuálnom prostredí. Takýmto 3R objektom je aj reálna ľudská hlava pre ktorú SNHC zavádza nové algoritmy kódovania založené na jej klonovaní pomocou geometrických modelov [5], animácie [6] ako aj textúrovania [7].

*Kľúčové slová* — MPEG-4, ľudská hlava, SNHC, modelovanie, animácia, textúrovanie, FAP, FDP

## I. ÚVOD

V súčasnosti informačné a komunikačné technológie sa rozširujú z reálneho prostredia aj do virtuálneho prostredia [1]. Vo všeobecnosti virtuálne prostredie (svet) umožňuje nahradiť reálne prostredie včítane jeho diania (života). To dáva možnosť rôznym geograficky vzdialeným účastníkom pôsobiť v tom istom virtuálnom prostredí za pomoci ich sieťového prepojenia. Potom v tomto prostredí môžu spoločne komunikovať, obchodovať, vzdelávať sa, zabávať sa, riešiť spoločné problémy atď. Takými virtuálnymi prostrediami môžu byť napr. virtuálne obchody, virtuálne úrady, virtuálne štúdiá, virtuálne knižnice, virtuálne múzea, virtuálne zábavné a vzdelávacie prostredia atď.

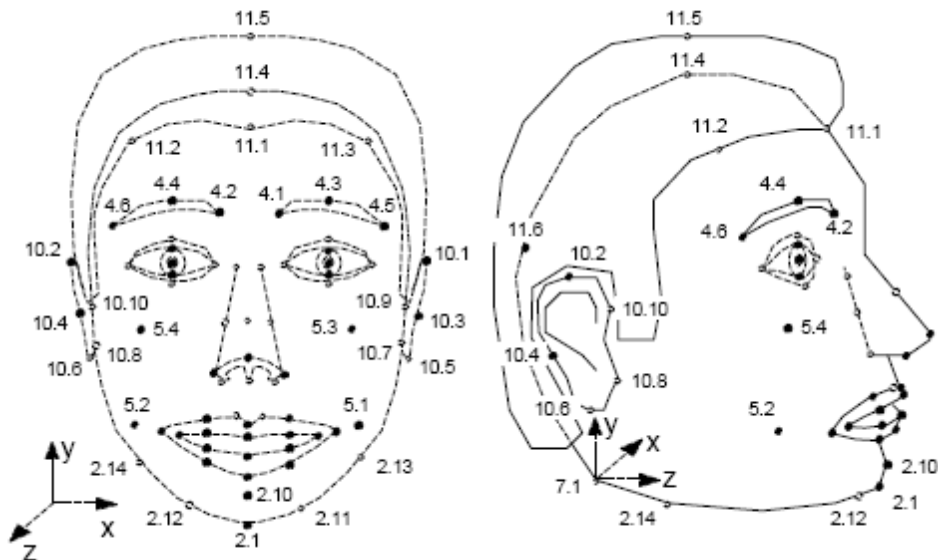
V tomto prostredí významným objektom je človek, ktorý v ňom môže byť reprezentovaný ako virtuál (avatar) alebo klon [2], [3]. Virtuál je jeho hrubou alebo symbolickou reprezentáciou bez toho, aby vyjadroval jeho reálnu podobu ako aj pohyb. Naopak klon predstavuje jeho vieruhodnú reálnu podobu včítane jeho pohybu. Vo všeobecnosti vo virtuálnom prostredí sa môžu spoločne vyskytovať ako virtuály tak aj klony. Pritom vedenie dialógu medzi nimi sa uskutočňuje pomocou ich ľudských hláv, ktoré môžu byť virtuálne alebo klonované. Potom virtuálne videokomunikácie umožňujú vzdialeným účastníkom vidieť tento dialóg v tom istom virtuálnom prostredí v ktorom sú nepriamo zúčastnení.

Vo všeobecnosti kódovanie ľudskej hlavy v štandarde MPEG-4 SNHC [8] používa jej analýzu výsledkom ktorej je množina parametrov pomocou ktorých sa vykoná syntéza klonovanej ľudskej hlavy [9]. Pritom táto analýza sa môže vykonávať v reálnom čase plne automaticky pre také aplikácie ako sú virtuálne videokonferencie. Naopak môže byť manuálne ovládaná nie v reálnom čase pre klonovanie ľudskej hlavy pre reprodukciu čítaného textu. Okrem toho napr. pre virtuálne obchodovanie nepotrebuje vykonať túto analýzu, pretože výsledkom syntézy nemusí byť klonovaná, ale len virtuálna hlava, ktorá nemusí napodobňovať reálnu hlavu.

## II. MODELOVANIE ĽUDSKEJ HLAVY

Analýza aj syntéza ľudskej hlavy v štandarde MPEG-4 SNHC používa jej geometrické modely z počítačovej grafiky. Pritom MPEG-4 SNHC špecifikuje pre tieto modely ich neutrálny (východzí) stav a zároveň maximálne 84 príznakových bodov ako to vidno na Obr.1. Tieto sú usporiadané do skupín vzťahujúce sa na určité časti ľudskej hlavy ako sú ústa, nos, oči atď. Pomocou príznakových bodov na reálnej ľudskej hlave možno potom tvarovať vybraný geometrický model vo východnom stave tak, aby jeho príznakové body boli v súlade s tými na reálnej ľudskej hlave. Súradnice príznakových bodov reálnej ľudskej hlavy predstavujú FDP (facial definition parameters) a v najjednoduchšom prípade ich možno priamo získať z výstupu 3R scanneru. Pretože 3R scanner sú málo dostupné a častejšie sa reálna ľudská hlava sníma

pomocou kamery, tieto FDP sa vypočítajú z dvoch ortogonálnych pohľadov, t.j. čelného a profilu tejto hlavy. Na základe znalosti FDP možno vykonať úpravu východzieho geometrického modelu pomocou DMS splajnovej interpolácie alebo FFD (free-form deformation), aby sa prispôbil jeho geometrický tvar k reálnej ľudskej hlave. Ďalšie metódy estimácie 3R súradníc príznakových bodov používajú sekvencie snímok reálnej ľudskej hlavy s rôznymi aj neortogonálnymi pohľadmi. Tieto umožňujú zmenšiť počet FDP, pretože pomocou interpolačných funkcií (radiálnych alebo splajnových) deformáciu geometrického modelu v mieste jeho korešpondujúcich príznakových bodov možno rozšíriť aj do ich okolia.



Obr. 1. Príklad zobrazenia príznakových bodov na čelnom a profilovom pohľade ľudskej hlavy.

### III. ANIMÁCIA GEOMETRICKÉHO MODELU ĽUDSKEJ HLAVY

Animácia umožňuje 3R globálny aj lokálny pohyb geometrického modelu ľudskej hlavy pomocou 68 animačných parametrov FAP (facial animation parameters) z ktorých 3 predstavujú parametre jeho globálnej rotácie a ostatné vyjadrujú malý translačný pohyb jednotlivých alebo malej skupiny vrcholov na jeho polygonálnej (trojuholníkovej) sieti. Pomocou tohto pohybu s následnou deformáciou okolia príslušných vrcholov možno interpretovať rôzne výrazy ľudskej tváre spôsobené lokálnym pohybom jej úst, očí, obočia, brady, uší, nosa atď. Preto je 68 animačných parametrov FAP zoradených do 10 skupín v závislosti od častí tváre ako to ukazuje Tabuľka 1.

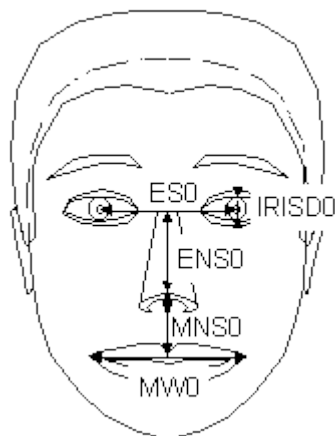
Tabuľka 1. FAP skupiny

Skupina	Počet FAPov
1. vizémy a výzory	2
2. čeľusť, brada, vnútro dolnej pery, štrbina medzi perami, kútiky	16
3. očné gule, zrenice, viečka	12
4. obočie	8
5. líca	4
6. jazyk	5
7. natočenie hlavy	3
8. pozície vonkajšej pery	10
9. nos	4
10. uši	4

Veľkosť tohto pohybu sa vyjadruje pomocou animačných jednotiek FAPU (facial animation parameter units) v Tabuľke 2, ktoré predstavujú zlomky charakteristických vzdialeností ľudskej tváre ako to vidno na Obr.2. Pritom tieto pohyby môžu byť unipolárne alebo bipolárne, kedy hodnoty animačných parametrov FAP sú z intervalu 0 až 1024 alebo -1024 až 1024. Smer týchto pohybov je určený animačným vektorom najčastejšie určený len jednou zložkou s hodnotou 1 pre jednotlivé vrcholy geometrického modelu ľudskej hlavy. Parametre globálnej rotácie sa vyjadrujú pomocou jednotky zlomku radiánu.

Tabuľka 2. Animačné jednotky FAPU

IRISD0	Priemer dúhovky (podľa definície je rovný vzdialenosti medzi horným a dolným viečkom) na modeli tváre v neutrálnom stave	$IRISD = IRISD0 / 1024$
ES0	Vzdialenosť očí	$ES = ES0 / 1024$
ENS0	Vzdialenosť očí a nosa	$ENS = ENS0 / 1024$
MNS0	Vzdialenosť úst a nosa	$MNS = MNS0 / 1024$
MW0	Šírka úst	$MW = MW0 / 1024$
AU	Jednotka uhla	10 exp.-5 rad



Obr. 2. Charakteristické vzdialenosti ľudskej tváre.

Prvú skupinu tvoria dva FAP parametre vyššej triedy. FAP 1 zodpovedá vizémam, ktoré sú vizuálne späté s fonémami. Je definovaných 14 statických vizém, ktoré sú zreteľne rozlíšiteľné ako je uvedené v Tabuľke 3. Aby sa dosiahol súlad reči a pohybov úst, nie je tvar úst hovoriaceho človeka ovplyvnený iba aktuálnou, ale aj predchádzajúcou a nasledujúcou fonémou. Prechody z jednej vizémi do druhej sú definované zmiešaním iba dvoch vizém s príslušnými váhovými faktormi. Pre vizuálnu animáciu reči s vysokou kvalitou bolo by potrebné zmiešavanie aj s väčším počtom vizém.

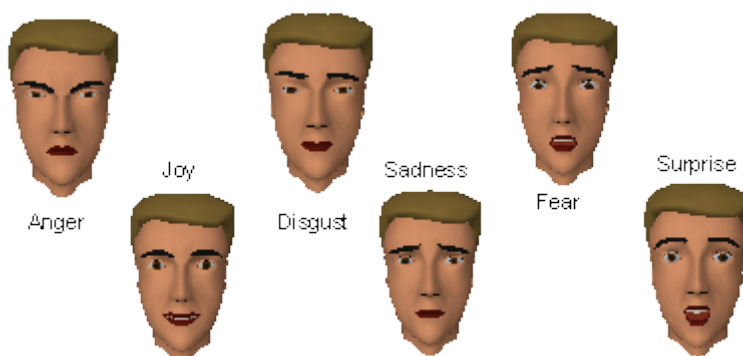
Tabuľka 3. Vizémy a príslušné fonémy

Vizéma #	fonémy	Príklad
0	nedefinovaná	–
1	p, b, m	put, <u>bed</u> , <u>mill</u>
2	f, v	<u>F</u> ar, <u>vo</u> ice
3	T, D	<u>th</u> ink, <u>th</u> at
4	t, d	<u>T</u> ip, <u>d</u> oll
5	k, g	<u>c</u> all, <u>g</u> as
6	tS, dZ, S	<u>ch</u> air, <u>jo</u> in, <u>sh</u> e
7	s, z	<u>S</u> ir, <u>z</u> eal
8	n, l	<u>L</u> ot, <u>n</u> ot
9	R	<u>R</u> ed
10	A:	<u>C</u> ar
11	E	<u>B</u> ed
12	I	<u>T</u> ip
13	Q	<u>T</u> op
14	U	<u>B</u> ook

FAP 2 definuje 6 základných výrazov ľudskej tváre ako to popisuje Tabuľka 4 a následne ilustruje Obr.3. Na rozdiel od vizém, výrazy tváre sú animované hodnotou FAP 2, ktorá definuje vybudenie výrazu. Dva výrazy ľudskej tváre môžu byť animované súčasne s amplitúdou v rozsahu  $<0;63>$  určenou pre každý výraz. Hodnoty parametra FAP 2 výzoru ľudskej tváre sú definované textovým popisom a konštruktér modelu ľudskej tváre ich určuje pre každý takýto model. Pretože sú určené pre celkový výraz ľudskej tváre, umožňujú animovať aj neznámy model s vysokou subjektívnou kvalitou.

Tabuľka 4. Základné výrazy ľudskej tváre definované pre FAP 2

#	Názov výrazu tváre	Textový popis
1	Radosť	Obočia sú v kľude. Ústa sú otvorené a kútiky úst stiahnuté späť, smerom k ušiam.
2	Smútok	Vnútoraná strana obočí je ohnutá dohora. Oči sú trochu privreté. Ústa sú v kľude.
3	Zlosť	Vnútoraná strana obočí je ohnutá dole a zmraštená. Oči sú široko otvorené. Pery sú stlačené proti sebe alebo otvorené aby odhalili zuby.
4	Strach	Obočia sú zdvihnuté a zmraštené. Vnútoraná strana obočia je ohnutá dohora. Oči sú napäté a v pohotovosti.
5	Odpor	Obočia a viečka sú uvoľnené. Horná pera je zdvihnutá a zvlhnená, často asymetricky.
6	Prekvapenie	Obočia sú zdvihnuté. Horné viečka sú široko otvorené, Dolné sú v kľude. Sánka je spadnutá.



Obr. 3. Základné výrazy ľudskej tváre definované pre FAP 2.

Používanie FAPov vyššej triedy (FAP 1, FAP 2) spolu s parametrami nižšej triedy (FAP 3-68 - pôsobia na rovnaké miesta ako FAP 1-2), môžu spôsobiť neočakávanú vizuálnu reprezentáciu ľudskej tváre. Všeobecne, FAPy nižšej triedy (3-68) majú prioritu pred deformáciami spôsobenými FAPmi vyššej triedy (1-2). Keď určujeme výraz tváre FAPom 2, môže kódér poslať „init-face“ bit, ktorý deformuje neutrálnu tvár modelu s výrazom pred použitými FAPmi 3-68. Táto deformácia je aplikovaná s vnúteným neutrálnym výrazom ľudskej tváre – zatvorenými ústami, otvorenými očami, smer pohľadu a orientácia hlavy. Pokiaľ kódér nevie nič o tom ako sú FAP 1 a 2 implementované, je odporúčané používať iba tie FAPy nižšej triedy, ktoré nebudú interferovať s FAP 1 a 2.

FAP môžu byť synteticky generované alebo sú výsledkom analýzy reálnej ľudskej hlavy. Pomocou extrakcie príznakov ako sú významné body, hrany, obrysy a ich následným sledovaním v snímkach videosekvencie reálnej ľudskej hlavy možno estimovať tieto parametre FAP. Pritom extrakciu príznakov možno robiť pomocou deformovateľných šablón jednotlivých častí ľudskej hlavy ako sú oči, ústa, atď alebo aplikáciu KLT na tieto časti. Metódy optického toku umožňujú vylúčiť potrebu extrakcie a sledovania príznakov v snímkach pre estimáciu FAP, pričom používajú parametrizovaný geometrický model alebo splajnovú aproximáciu povrchu



ľudskej hlavy [10]. Metódy založené na AAM (active appearance model) nepočítajú optický tok, ale pomocou PCA (principal component analysis) estimujú špecifické parametre globálneho a lokálneho pohybu parametrizovaného geometrického modelu, ktoré sa následne konvertujú na FAP [11].

#### IV. ZÁVER

Vo všeobecnosti výsledkom analýzy reálnej ľudskej hlavy po jej detekcii a lokalizácii vo vstupnej videosekvencii na vysielačnej strane sú parametre FDP, FAP a textúry, pričom FDP a textúry sa vysielaajú iba v prvej snímke pre syntézu klonovanej alebo sa nevysielaajú vôbec pre syntézu virtuálnej ľudskej hlavy vo virtuálnom prostredí. Potom pravidelne po snímkach sa vysielaajú iba parametre FAP čo vedie k veľmi nízkym bitovým rýchlostiam v rozsahu 300 až 2000 bit/s. FAP možno kódovať s menšou kompresiou aj oneskorením pomocou medzisnímkových diferenčných metód s aritmetickým kódovaním [12]. Naopak väčšia kompresia, ale aj oneskorenie sa dosahuje pomocou medzisnímkového transformačného kódovania s DCT [13]. Pretože FDP sa vysielaajú iba v prvej snímke na ich kódovanie možno jednoducho použiť PCM. Textúry sa kódujú podobne ako statické obrazy pomocou vnútrosnímkových transformačných metód v ktorých sa najviac používa diskretná waveletová transformácia (DWT) [14].

#### LITERATÚRA

- [1] Ch. Joslin - I. S. Pandzic - N.M. Thalmann, "Trends in Networked Collaborative Virtual Environments," *Computer Communications*, vol. 26, no. 5, 2003, pp. 430-437.
- [2] N. M. Thalmann – D. Thalmann, *Handbook of Virtual Humans*. John Wiley & Sons, 2004.
- [3] O. N. N. Fernando - G. Saito - U. Duminduardena - Y. Tanno - M.Cohen, Cloning and Teleporting Avatars Across Workstations and Mobile Devices in Collaborative Virtual Environments: Clipboard Operations for Virtual Reality. *Proc. Intl. Conf. on Information and Automation*. Colombo, Sri Lanka, 2005, pp. 133-140.
- [4] J. Mihalík, "Štandardný videokodek MPEG-4," *Slaboproudý obzor*, roč. 60, č.2, 2003, s. 7-11.
- [5] J. Mihalík, "Modeling of Human Head Surface by using Triangular B-Splines," *Radioengineering*, vol. 19, no. 1, 2010, pp. 39-45.
- [6] J. Mihalík - V. Michalčín, "Animation of 3D Model of Human Head," *Radioengineering*, vol. 16, no. 1, 2007, pp. 58-66.
- [7] J. Mihalík - V. Michalčín, "Texturing of Surface of 3D Human Head Model," *Radioengineering*, vol. 13, no. 4, 2004, pp. 44-47.
- [8] *The special issue of the IEEE Trans.on Circuits and Systems for Video Technology on MPEG-4 SNHC*, July 2004.
- [9] J. Mihalík, "3R číslicové spracovanie videoobjektu ľudskej hlavy," LDIPV, FEI TU Košice, 2019.
- [10] J. Mihalík, "Číslicové spracovanie obrazov," LDIPV, FEI TU Košice, 2018.
- [11] J. Mihalík - M. Kasár, "Basis of Eigenfaces for Tracking of Human Head," *Journal of Electrical Engineering*, vol. 58, no. 3, 2007, pp. 134-139.
- [12] I. Gladišová - J. Mihalík, "Stavové aritmetické kódovanie binárnych obrazov," *Acta Electrotechnica et Informatica*, vol. 3, No. 3, 2003, pp. 36-44.
- [13] J. Mihalík, "Adaptívne transformačné kódovanie obrazu," *Slaboproudý obzor*, 1991, roč. 52, č.11-12, s. 253-257.
- [14] J. Zavacký - J. Mihalík - I. Gladišová, "Implementácia diskretnéj waveletovej transformácie v štandarde JPEG-2000," *Slaboproudý obzor*, roč. 63, č.3-4, 2007, s. 5-9.

# UWB-PerLoc-2D3D: softvér na lokalizáciu osôb v 2D a 3D priestore

<sup>1</sup>Michal ŠVINGÁL, <sup>2</sup>Dušan KOCUR, <sup>3</sup>Jana FORTES

<sup>1</sup>K-Mlab organizačná zložka Imsens GmbH, Košice, Slovensko

<sup>2,3</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

<sup>1</sup>Michal.Svingal@ilmsens.com, <sup>2</sup>Dusan.Kocur@tuke.sk, <sup>3</sup>Jana.Fortes@tuke.sk

**Abstrakt** – Lokalizácia pohybujúcich sa osôb v 2D alebo 3D priestore pomocou UWB (z angl. *Ultra-WideBand*) senzorov krátkého dosahu je v celosvetovom meradle predmetom intenzívneho výskumu. V predložennom príspevku opisujeme návrh nového softvéru nazvaného UWB-PerLoc-2D3D. Tento softvér vytvorený v programovacom prostredí MATLAB je určený na spracovanie UWB radarových signálov s cieľom detegovať, lokalizovať a sledovať pohybujúce sa osoby v 2D a 3D priestore. V závere príspevku uvádzame zhodnotenie vlastností navrhnutých programových prostriedkov a ich využitie v praxi.

**KLúčové slová** – lokalizácia osôb, softvér v programovacom prostredí MATLAB, spracovanie radarových signálov, UWB radar

## I. ÚVOD

Pri riešení mimoriadnych situácií, kedy môže dochádzať k ohrozeniu ľudských životov, je možné využiť celý rad praktických aplikácií lokalizácie osôb pomocou UWB (z angl. *Ultra-WideBand*) radarov krátkého dosahu. Použitie týchto aplikácií je výhodné v prípadoch, kedy nie je možné využiť štandardné prostriedky určené na vyhľadávanie postihnutých osôb. Ide o situácie, v ktorých sa zistí uje poloha osôb pri nízkej až nulovej optickej viditeľnosti. V tomto prípade ide napríklad o osoby nachádzajúce sa v prostredí s veľkou koncentráciou prachu, dymu alebo osoby zavalené sutinami budov.

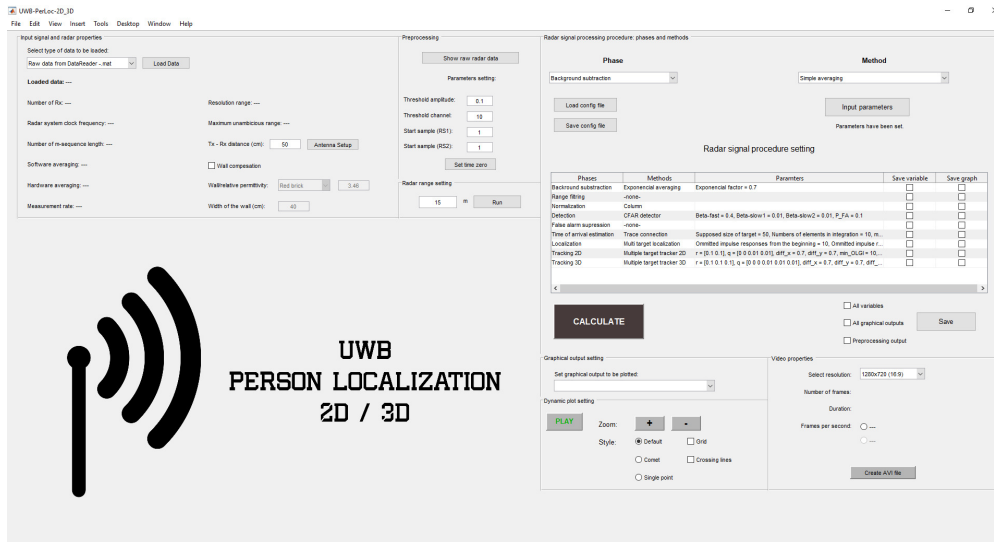
V súčasnej dobe je dôležité zabezpečiť kvalitný život ľuďom vo vyššom veku. Mnohí ľudia nemôžu z rôznych príčin zabezpečiť seniorom dennú starostlivosť. Často sa stáva, že starší človek si nepraje byť monitorovaný kamerovým systémom. Možnosť asistovaného žitia pomocou UWB senzorov je jednou z alternatív detekcie ich neštandardného správania sa. UWB radarový systém je využívaný na monitorovanie životných funkcií osôb a zmeny ich telesnej polohy (napr. detekcia netypickej polohy).

Oblasť spracovania údajov nameraných UWB senzorom je predmetom nepretržitého a intenzívneho výskumu. Za účelom vývoja aplikácií na lokalizáciu osôb UWB senzormi je praktické disponovať programovými prostriedkami, ktoré umožňujú spracovávať namerané radarové signály. Na Katedre elektroniky a multimediálnych telekomunikácií Technickej univerzity v Košiciach (TUKE) je uskutočňovaný výskum v oblasti UWB lokalizačných systémov, ktorého súčasťou je aj vývoj softvéru aplikovateľného na výpočet metód spracovania signálov nameraných UWB senzormi. Jednou z úloh tohto softvéru je prostredníctvom získaných radarových signálov detegovať, lokalizovať a sledovať pohybujúce sa osoby [1].

V rámci diplomovej práce [2] bol v programovacom prostredí MATLAB vytvorený nový softvér nazvaný UWB-PerLoc-2D3D. Je určený na off-line spracovanie nameraných UWB radarových signálov s cieľom lokalizovať pohybujúce sa osoby v 2D a 3D priestore. Ako predloha pre návrh tohto programu bol použitý program UWB-PerLoc-MAT, ktorý bol vyvinutý v rámci bakalárskej práce [3]. Na rozdiel od tohto softvéru UWB-PerLoc-2D3D disponuje mnohými vylepšeniami, ktoré sú opísané v nasledujúcich častiach tohto príspevku.

## II. NÁVRH SOFTVÉRU UWB-PERLOC-2D3D

Účelom softvéru UWB-PerLoc-2D3D v programovacom prostredí MATLAB je zabezpečiť efektívne nastavenie parametrov jednotlivých metód procedúry spracovania UWB radarových signálov na lokalizáciu pohybujúcich sa osôb (v ďalšom procedúra SPMP, z angl. *Signal processing Procedure for Moving Persons*) a následne spracovať signály namerané pomocou UWB senzorov



Obr. 1 Vzhľad grafického rozhrania softvéru UWB-PerLoc-2D3D.

so zámerom lokalizovať pohybujúce sa osoby v 2D alebo v 3D priestore [4], [5]. Jedná sa teda o systém off-line spracovania UWB radarových signálov.

Pri vývoji programu UWB-PerLoc-2D3D sme vychádzali z bakalárskej práce [3], v rámci ktorej bol vytvorený softvér UWB-PerLoc-MAT určený na spracovanie signálov získaných z radarového systému pracujúceho s 1 Tx a 2 Rx. Na rozdiel od tohto softvéru je program UWB-PerLoc-2D3D rozšírený o algoritmy lokalizácie osôb v 3D priestore [6], [7]. Na základe toho bolo nutné upraviť reťazec zabezpečujúci výpočet SPMP tak, aby vykonával odhad vzdialeností pohybujúcich sa osôb z radarových dát získaných zo štyroch prijímacích kanálov. Taktiež novovytvorený softvér oproti predošlému disponuje mnohými vylepšeniami a novými funkciami z hľadiska ovládateľnosti softvéru a aj programovej implementácie. Tie najdôležitejšie sú stručne opísané v nasledujúcich podsekcích.

#### A. Návrh štruktúry priečinkov pre lepšiu orientáciu medzi jednotlivými m-súbormi

V pôvodnej verzii softvéru sa všetky potrebné m-súbory nachádzali v jednom priečinku. Takáto štruktúra súborov sa javí nepraktická zo strany vývojára. Preto bola navrhnutá nová štruktúra, ktorá zabezpečuje to, že m-súbory sprostredkujúce výpočet jednotlivých fáz SPMP sa nachádzajú v samostatných priečinkoch, čo v konečnom dôsledku zlepšuje orientáciu medzi algoritmami softvéru [2].

#### B. Použitie koncepcie „vnorených funkcií“ (z angl. nested functions) pre programovanie hlavnej funkcie programu

Keďže navrhovaný softvér disponuje užívateľským grafickým rozhraním (GUI, z angl. *graphical user interface*), bolo potrebné toto rozhranie v rámci vývoja softvéru definovať. V programe UWB-PerLoc-MAT bol na tento účel použitý nástroj GUIDE, ktorého výhody sú detailne opísané v [3]. Ukázalo sa však, že pre návrh niektorých ďalších funkcií softvéru je vhodnejšie mať grafické rozhranie definované programovo použitím funkcií MATLABU. Preto pri programovaní hlavnej funkcie softvéru bola zvolená koncepcia vnorených funkcií [8], ktorej použitie prinieslo nasledujúce výhody:

- odstránenie globálnych premenných,
- možnosť nastavovania spôsobu zobrazovania jednotlivých interaktívnych komponentov pri zmene veľkosti hlavného okna aplikácie,
- možnosť implementácie panelu nástrojov, ktorý slúži na zmenu spôsobu zobrazovania grafických výstupov.

Grafické rozhranie programu UWB-PerLoc-2D3D je zobrazené na Obr. 1.

#### C. Možnosť načítania dát rôzneho typu, resp. formátu

Softvér UWB-PerLoc-MAT sa vyznačuje nejasnosťou zadávania typu načítaných dát. Táto nejasnosť vyplýva z toho, že pri výbere typu dát, ktoré majú byť softvérom načítané, sú na výber možnosti „Radar 1“, „Radar 2“ a „Radar 3“ podľa toho, s ktorým UWB senzorom bolo

vykonané meranie. Omnoho lepšie riešenie výberu typu dát spočíva v uvedení typov real-time softvérov, s ktorými bolo uskutočnené meranie, a im prislúchajúcim formátom výstupných dát. Taktiež bola implementovaná aj voľba týchto typov dát: pedspracované radarové dáta, dáta s odčítaným pozadím, filtrované dáta, normované dáta, výstup detektora, výstup fázy potlačenia falošných cieľov.

#### *D. Implementácia ďalších algoritmov spracovania radarových signálov*

Jednou z najväčších nevýhod pôvodného softvéru je počet implementovaných metód SPMP. Zdrojom programových implementácií týchto metód bol softvér Toolbox\_UWB, ktorý bol navrhnutý na TUKE a slúži na spracovanie radarových signálov s cieľom lokalizovať pohybujúce sa osoby v 2D priestore. Podrobnejšie je tento softvér opísaný v [3]. Z dôvodu čo najlepšej efektívnosti lokalizácie osôb boli do UWB-PerLoc-MAT implementované len tie algoritmy, ktoré poskytujú najlepšie výsledky. Pri vývoji nových algoritmov je však vhodné disponovať aj ostatnými metódami s cieľom analýzy ich funkčnosti pre nasledujúci výskum. Práve preto bola do UWB-PerLoc-2D3D implementovaná celá rada ďalších algoritmov SPMP, z ktorej spomenieme len algoritmus kompenzácie efektu steny. Absencia tohto algoritmu v pôvodnom softvéri spôsobovala značné nepresnosti lokalizácie v prípade, ak boli spracovávané dáta z meraní cez stenu. Preto obsahuje novo vytvorený softvér aj možnosť nastavenia parametrov steny [9].

#### *E. Efektívnejšia realizácia programovej implementácie hlavného reťazca operácií aplikovaného na výpočet SPMP*

Jednou z úloh diplomovej práce [2] bolo vytvorenie toolboxu v programovacom prostredí MATLAB, ktorý obsahuje samotný program UWB-PerLoc-2D3D, ale aj jednotlivé algoritmy metód spracovania signálov pre možnosť vytvárania ďalších programových prostriedkov realizujúcich výpočet SPMP. Pri príprave týchto algoritmov bola vytvorená funkcia *UWB\_processing*, ktorú možno efektívne využiť na naprogramovanie celého reťazca operácií slúžiaceho na výpočet SPMP. Takýto postup bol aplikovaný aj v softvéri UWB-PerLoc-2D3D. Oproti pôvodnému softvéru je časová náročnosť výpočtu SPMP omnoho menšia, nakoľko program počas realizácie výpočtu jednotlivých metód nevykonáva zobrazenie každého grafického výstupu reťazca operácií, ale zobrazí len grafický výstup poslednej zvolenej metódy.

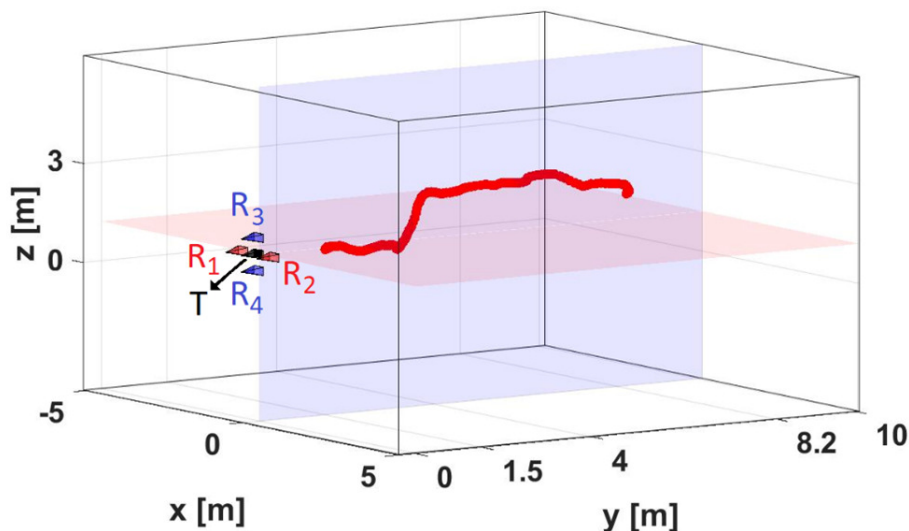
#### *F. Oprava procesu ukladania grafických výstupov*

Pri ukladaní grafických výstupov bolo v pôvodnej verzii softvéru nutné zobrazit každý grafický výstup v samostatnom okne ešte pred tým, ako program vykonal uloženie tohto výstupu. Následne po uložení sa zobrazené okno zatvorilo. Z tohto dôvodu bolo ukladanie väčšieho množstva grafických výstupov veľmi nepraktické a z užívateľskej stránky to pôsobilo ako chyba softvéru. Táto nedokonalosť bola v novej verzii potlačená. Taktiež je v novo vytvorenej verzii softvéru pridaná možnosť uloženia grafických výstupov vo formátoch .jpg, .png, .fig a .avi.

### III. VLASTNOSTI SOFTVÉRU UWB-PERLOC-2D3D

Najzásadnejším zdokonalením softvéru UWB-PerLoc-2D3D oproti pôvodnému softvéru UWB-PerLoc-MAT, okrem vyššie opísaných vylepšení z hľadiska ovládateľnosti softvéru a programovej implementácie, je schopnosť spracovať UWB radarové signály získané z multistatického radaru pracujúceho s 1 Tx a 4 Rx, a teda lokalizovať pohybujúce sa osoby v 3D priestore. Táto vlastnosť bola zabezpečená implementáciou nových lokalizačných algoritmov. Ide o 3D-2D metódu, aproximáciu 3D-2D metódy, optimalizačnú metódu a Newtonovu iteračnú metódu [5], [6], [7]. Vzhľadom na to, že týmto algoritmom z hľadiska správnosti programovej implementácie nebola venovaná dostatočná pozornosť, bola vykonaná analýza ich vlastností na vzorke syntetických dát pomocou počítačovej simulácie. Z dôvodu prípravy funkcií zabezpečujúcich výpočet uvedených algoritmov a funkcií pre zobrazovanie odhadnutých súradníc cieľov v 3D súradnicovej sústave sa ukázalo vhodné vytvoriť túto simuláciu v programovacom prostredí MATLAB. Hlavným prínosom simulácie bola príprava rôznych ďalších funkcií určených na výpočet nielen vyššie uvedených algoritmov, ale aj celej rady ďalších algoritmov, ktoré boli vytvorené ako kombinácie dostupných algoritmov lokalizácie a sledovania cieľov v 2D a 3D priestore. Keďže navrhnutá simulácia je založená na testovaní lokalizačných algoritmov na vzorke syntetických dát, bolo možné analyzovať tieto algoritmy aj z hľadiska presnosti lokalizácie. Uvedený postup zásadne uľahčil návrh programu UWB-PerLoc-2D3D, a preto možno konštatovať, že tento postup bol vhodný pre riešenie úlohy implementácie daných algoritmov do novo vytváraného softvéru.

V rámci diplomovej práce [2] bol taktiež navrhnutý toolbox UWB Toolbox v programovacom prostredí MATLAB, ktorý obsahuje funkcie sprostredkujúce spracovanie nameraných UWB



Obr. 2 Grafický výstup 3D lokalizačného algoritmu aplikovaného na experimentálne dáta.

radarových signálov. Tieto funkcie sú vytvorené tak, aby ich bolo možné využiť na vytvorenie ďalších možných programových prostriedkov zameraných na vývoj aplikácií UWB senzorov. Pri vývoji týchto funkcií bola definovaná základná koncepcia výpočtu procedúry spracovania radarových signálov, ktorá bola použitá aj v programe UWB-PerLoc-2D3D, a taktiež bola použitá v jednoduchých programoch vytvorených za účelom demonštrácie jej použitia pre budúcich vývojárov.

S cieľom overenia vlastností softvéru UWB-PerLoc-2D3D bolo v rámci diplomovej práce [2] uskutočnené experimentálne meranie s multistatickým UWB radarom a následne boli získané dáta spracované pomocou navrhnutého programu. Zo spracovania týchto dát možno tvrdiť, že softvér je schopný lokalizovať pohybujúce sa osoby v 3D priestore, a teda je vytvorený podľa špecifikácie zadania (Obr. 2). Okrem výpočtu celej procedúry spracovania radarových signálov však softvér disponuje množstvom ďalších pozitívnych vlastností, medzi ktoré patrí dynamické vykresľovanie grafických výstupov, možnosť načítania dát získaných jedným z dvoch typov real-time softvérov a aj spracovaných dát z niektorých fáz procedúry, ukladanie grafických výstupov programu vo formátoch .jpg, .png, .fig, alebo aj ukladanie videa vo formáte .avi.

Samozrejme, navrhnuté programové prostriedky sa vyznačujú aj negatívnymi vlastnosťami. V prípade UWB-PerLoc-2D3D je jeho najväčšou nevýhodou veľká pamäťová náročnosť. Tento program je vytvorený tak, že všetky premenné obsahujúce spracované signály sú po výpočte procedúry stále dostupné v pamäti pre potreby ich zobrazenia, alebo akejkoľvek inej operácie. Túto vlastnosť by bolo možné eliminovať použitím iného prístupu programovania, pri ktorom program pracuje s úložiskom, resp. pri výpočte danej premennej program túto premennú uloží na vopred definované miesto na úložisku. Následne pri požiadavke vykonania nejakej operácie, napríklad zobrazenia grafického výstupu, je táto premenná načítaná do pamäti. V prípade funkcie *UWB\_processing* toolboxu UWB Toolbox je jej negatívnou vlastnosťou pomerne zložitý proces definovania vstupných parametrov a taktiež prístupovania k výstupným parametrom, nakoľko vstupy a výstupy tejto funkcie sú definované ako polia typu *cell* [10]. Táto zložitosť vyplýva z toho, že pri návrhu tejto funkcie bolo cieľom vytvoriť koncepciu, podľa ktorej možno jednoducho naprogramovať celú procedúru spracovania UWB radarových signálov.

#### IV. ZÁVER

Zo spôsobu práce s programom UWB-PerLoc-2D3D, ktorá je vďaka grafickému rozhraniu jednoduchá a intuitívna, možno konštatovať, že tento softvér je možné využívať vo vzdelávacej a vedeckovýskumnej činnosti na TUKE. Jeho pilotné testovanie prebehlo v rámci predmetu UWB senzorové siete v zimnom semestri školského roku 2020/2021. Softvér UWB-PerLoc-2D3D je úspešne používaný aj v rámci výskumu a vývoja v oblasti monitorovania osôb uskutočňovaného na Katedre elektroniky a multimediálnych telekomunikácií.

Vzhľadom na to, že v mnohých prípadoch sa ukazuje praktické disponovať softvérom, ktorý možno použiť na lokalizáciu pohybujúcich sa osôb v reálnom čase, námetom na rozšírenie tejto práce je práve implementácia takejto funkcionality do softvéru UWB-PerLoc-2D3D.

## POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe projektu č. APVV-18-0373 a vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a Slovenskej akadémie vied (VEGA) na základe projektu č. 1/0584/20.

## LITERATÚRA

- [1] Kocur, D. et al., *Detekcia, lokalizácia a sledovanie osôb UWB radarmi krátkeho dosahu*. TUKE, Košice, 2016. ISBN 987-80-553-2648-1
- [2] Švingál, M., *UWB-PerLoc-2D3D: Softvér určený na vývoj aplikácií UWB senzorov na lokalizáciu osôb v 2D a 3D priestore*. Diplomová práca, TUKE, Košice, 2019.
- [3] Švingál, M., *UWB-PerLoc-MAT: Softvér v programovacom prostredí MATLAB určený na vývoj aplikácií UWB senzorov na lokalizáciu osôb*. Bakalárska práca, TUKE, Košice, 2017.
- [4] Rovňáková, J., *Complete signal processing for through wall target tracking by M-sequence UWB radar system*. Dizertačná práca, TUKE, Košice, 2009.
- [5] Švecová, M., *Lokalizácia cieľa UWB radarovým systémom*. Dizertačná práca, TUKE, Košice, 2009.
- [6] Švecová, M., Kocur, D., Demčák, J., Gamcová, M., Slovák, S. "Through-the-floor localization of a static person by multistatic UWB radar." *Microwave and Optical Technology Letters*, 1–7, 2018.
- [7] Buša, J., "Target Localization by UWB Signals," In: Dimov I., Fidanova S., Lirkov I. (eds) *Numerical Methods and Applications*. NMA 2014. Lecture Notes in Computer Science, vol 8962, s. 254–261, 2015. *Springer, Cham*, DOI : 10.1007/978 – 3 – 319 – 15585 – 228.
- [8] MathWorks: Nested Functions. [online]. The MathWorks, Inc. [citované 13.5.2021]. Dostupné na internete < [https://www.mathworks.com/help/matlab/matlab\\_prog/nested\\_functions.html](https://www.mathworks.com/help/matlab/matlab_prog/nested_functions.html) >.
- [9] Rovňáková, J., Kocur, D., "Compensation of wall effect for through wall tracking of moving targets," In: *Radioengineering, Part 2, Special Issue on Workshop of the COST Action IC0803: RF/Microwave Communication Subsystems for Emerging Wireless Technologies (RFCSET)*. Vol. 18, no. 2, 2009, p. 189-195. ISSN 1210-2512.
- [10] MathWorks: Cell Arrays. [online]. The MathWorks, Inc. [citované 13.5.2021]. Dostupné na internete < [https://www.mathworks.com/help/matlab/cell\\_arrays.html](https://www.mathworks.com/help/matlab/cell_arrays.html) >.

# Vylepšenie vlastností logistického zobrazenia z hľadiska obrazových šifrovacích algoritmov

Jakub ORAVEC, Ľuboš OVSENÍK, Maroš LAPČÁK,  
Norbert ZDRAVECKÝ

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

{jakub.oravec, lubos.ovsenik, maros.lapcak, norbert.zdravecky}@tuke.sk

**Abstrakt** – V tomto príspevku sa budeme zaoberať viacerými technikami, ktoré slúžia na vylepšenie vlastností logistického zobrazenia. Aj keď je toto zobrazenie pomerne jednoduché, v niektorých špecifických aplikáciách je stále pomerne často využívané. Ku takýmto použitiam patria napríklad aj dedikované obrazové šifrovacie algoritmy, pri ktorých je potrebné klásť dôraz na obmedzenie nevýhod logistického zobrazenia. Najznámejšími z nich sú prílišný vplyv hodnoty parametra zobrazenia na vlastnosti generovaných postupností, závislosť medzi susednými prvkami postupnosti, či existencia pevného bodu zobrazenia. Predstavené techniky by mali dopomôcť k potlačeniu týchto nevýhod, na čo poukazujú aj jednoduché príklady dopadov navrhnutých riešení vo forme obrázkov, respektíve tabuľky.

**Keľúčovú slová** – Ljapunovov exponent, logistické zobrazenie, obrazové šifrovacie algoritmy, teória chaosu

## I. ÚVOD

Logistické zobrazenie (LZ) predstavovalo v minulosti relatívne populárny nástroj na vytvorenie generátorov pseudo-náhodných (PN) postupností [1]–[3]. Neskôr však niekoľko publikácií poukázalo na rôzne nevyhovujúce vlastnosti LZ [4], [5]. Časť z týchto vlastností bola odstránená pri návrhu nových obrazových šifrovacích algoritmov [6]–[8]. Tieto algoritmy boli preskúmané technikami kryptoanalýzy a zo zistených vlastností algoritmov sa vyvodili všeobecné závery [9], [10]. Tieto závery ukazujú na skutočnosť, že pri správnej implementácii a dodržaní niekoľkých pravidiel je možné použiť aj LZ [11].

V našich predošlých publikáciách sme sa zaoberali možnosťami pozmenenia predpisu LZ [12], resp. rôznymi technikami, ktoré vylepšujú vlastnosti generovaných PN postupností [13], [14]. Získané poznatky plne korešpondujú s aktuálnym stavom poznania tejto problematiky. V tomto príspevku bude predstavených niekoľko techník, ktoré by mali vylepšiť vlastnosti LZ.

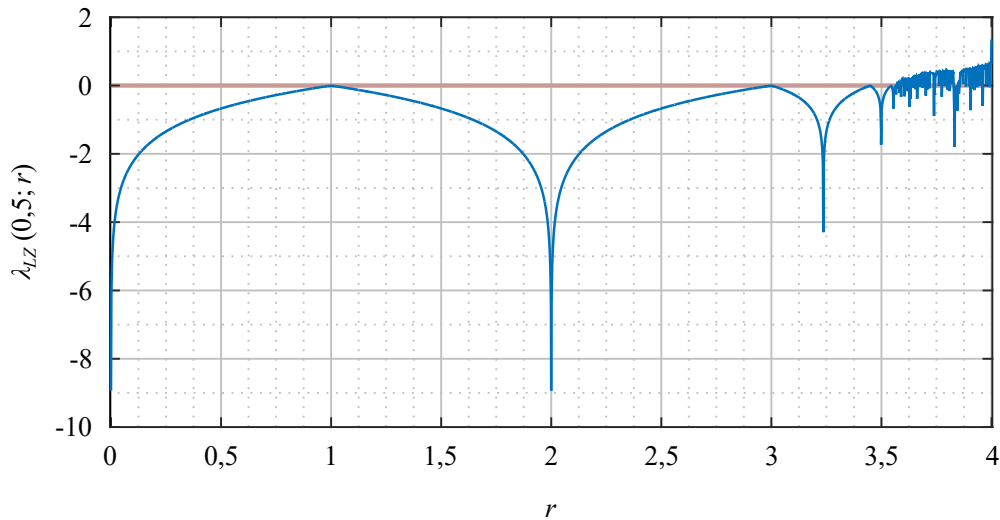
## II. LOGISTICKÉ ZOBRAZENIE A JEHO VLASTNOSTI

LZ je jednorozmerné chaotické zobrazenie, ktoré využíva parameter  $r \in (0; 4)$  a zobrazuje hodnotu iterátu  $x_n \in (0; 1)$  na hodnotu nasledujúceho iterátu  $x_{n+1} \in (0; 1)$ . Hodnota iterátu  $x_0$  sa označuje aj ako počiatočná hodnota a zvyčajne je zvolená ako 0,5 [1], [3]. Matematický predpis LZ je možné vyjadriť rovnicou (1):

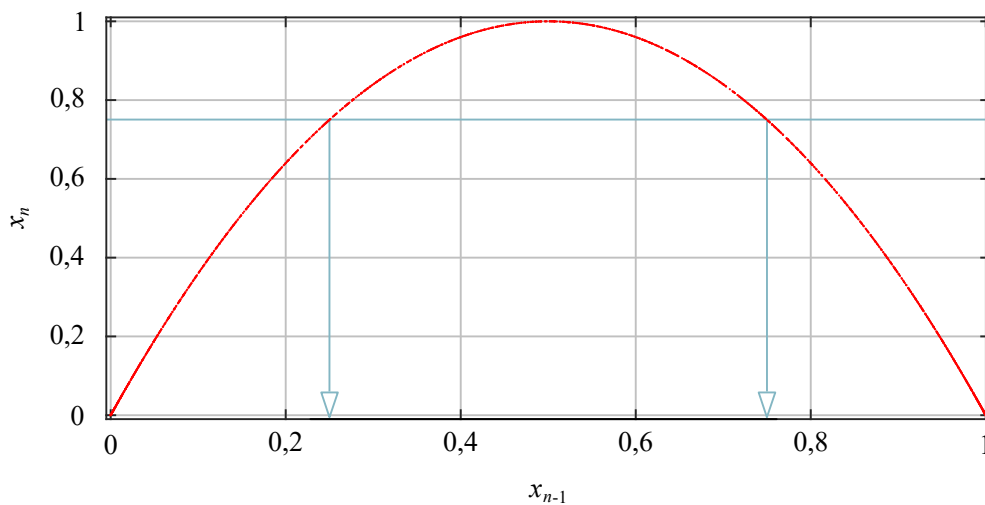
$$x_{n+1} = r \cdot x_n(1 - x_n). \quad (1)$$

Pri predpise LZ treba spomenúť aj existenciu prechodnej doby – určitá časť iterátov sa nepoužíva priamo na generovanie prvkov PN postupností, ale iba na zmenu hodnoty, ktorá sa použije namiesto počiatočnej hodnoty  $x_0$ . Táto technika a jej efekt sú dostatočne popísané vo viacerých prameňoch [1], [3], [12], [14].

Chaotické vlastnosti LZ, ktoré sú žiadané z hľadiska obrazových šifrovacích algoritmov je možné kvantifikovať pomocou hodnoty Ljapunovovho exponentu  $\lambda$ . Pre prípad LZ je možné hodnotu  $\lambda_{LZ}$  odhadnúť vzťahom (2) [3], [11]:



Obr. 1 Priebeh hodnôt Ljapunovovho exponentu pre LZ



Obr. 2 Poincarého diagram pre LZ s hodnotou parametra  $r = 4 - 10^{-15}$

$$\lambda_{LZ} \sim \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=0}^{m-1} \ln |r(1 - 2x_n)|, \quad (2)$$

kde zápis  $\ln a$  vyjadruje prirodzený logaritmus čísla  $a$  a zátvorky  $|b|$  označujú absolútnu hodnotu čísla  $b$ . Najčastejšie používanou hodnotou parametra  $m$  je 1000 alebo 10000 iterácií [11].

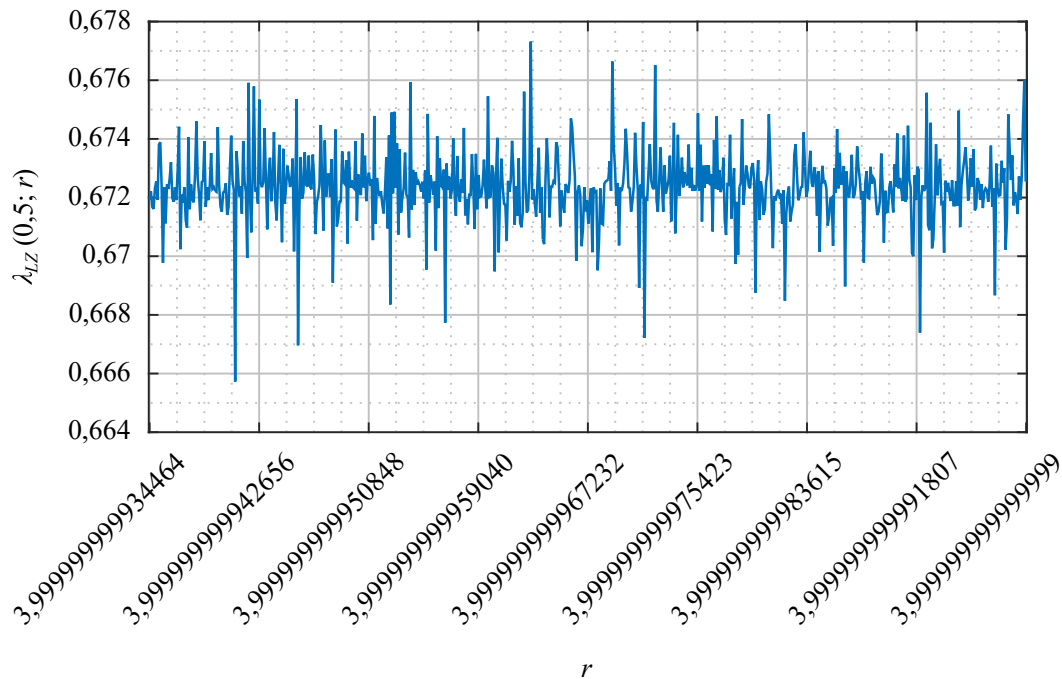
Zo vzťahu (2) vyplýva, že hodnota  $\lambda_{LZ}$  je funkciou použitej hodnoty parametra  $r$ . Na túto skutočnosť poukazuje aj Obr. 1, kde sa pomocou vzťahu (2) vypočítali hodnoty  $\lambda_{LZ}$  pre  $x_0 = 0,5$  a  $r \in (0; 4)$ .

Hodnoty  $\lambda_{LZ}$  je možné interpretovať jednoduchým spôsobom – ak  $\lambda_{LZ} > 0$ , tak LZ pri zvolenej hodnote parametra  $r$  nadobúda chaotické vlastnosti. V opačnom prípade je možné vlastnosti postupnosti iterátov s určitou dĺžkou pomerne jednoducho predpovedať [4], [5].

Okrem obmedzeného intervalu hodnôt parametra  $r$ , ktorý je možné použiť na generovanie PN postupností je potrebné poukázať aj na skutočnosť, že podľa (1) je hodnota  $x_{n+1}$  funkciou hodnoty  $x_n$ . Graficky to ilustruje aj Poincarého diagram zobrazený na Obr. 2. Pri použití hodnôt  $x_n = 0,25$  a  $r = 4 - 10^{-15}$  získame v dátovom type double precision [15] hodnotu  $x_{n+1} = 0,75$ . Kolmice na úsečku, ktorá má na vertikálnej osi súradnicu rovnú hodnote  $x_{n+1}$  v miestach, kde pretína charakteristickú parabolu LZ ukazujú možné hodnoty  $x_n$ .

Ďalším z identifikovaných problémov LZ je existencia tzv. pevného bodu [4], [5]. Ak sa hodnota iterátu rovná hodnote pevného bodu, všetky nasledujúce iteráty budú mať rovnakú hodnotu. Hodnota pevného bodu závisí od parametra  $r$ , dá sa vypočítať ako  $\frac{1}{1-r}$  [1], [3], [11].





Obr. 3 Priebeh hodnôt Ljapunovovho exponentu pre vybraný interval hodnôt parametra  $r$

### III. NAVRHNUTÉ RIEŠENIA

V našich predošlých publikáciách boli predstavené riešenia pre vyššie uvedené problémy [11]–[14]. Kvôli obmedzenej dĺžke tohto príspevku sa zameriame iba na ich krátky opis.

#### A. Zabezpečenie vhodných vlastností LZ

Požadované chaotické vlastnosti LZ je možné doceliť pri použití hodnôt parametra  $r$ , ktoré dosahujú čo najvyššie hodnoty  $\lambda_{LZ}$ . Ak uvažujeme o chaotickom šifrovacom algoritme, v ktorom použitý kľúč priamo ovplyvní hodnotu parametra  $r$ , môžeme uvažovať o intervale  $r \in \langle 4 - 65536 \cdot 10^{-15}; 4 - 10^{-15} \rangle$ . Takto získame v dátovom type double precision 65536 rôznych hodnôt  $r$ , čo zodpovedá časti kľúča s dĺžkou  $\log_2 65536 = 16$  bitov.

Pri použití väčších intervalov hodnôt  $r$ , ktoré by umožnili použitie dlhších častí kľúča boli zistené výraznejšie rozdiely v priebehu hodnôt  $\lambda_{LE}$ . Túto problematiku sme detailne opísali v [11]. Hodnoty  $\lambda_{LE}$  pre  $r \in \langle 4 - 65536 \cdot 10^{-15}; 4 - 10^{-15} \rangle$  ukazuje Obr. 3.

#### B. Potlačenie závislosti medzi susednými iterátmi

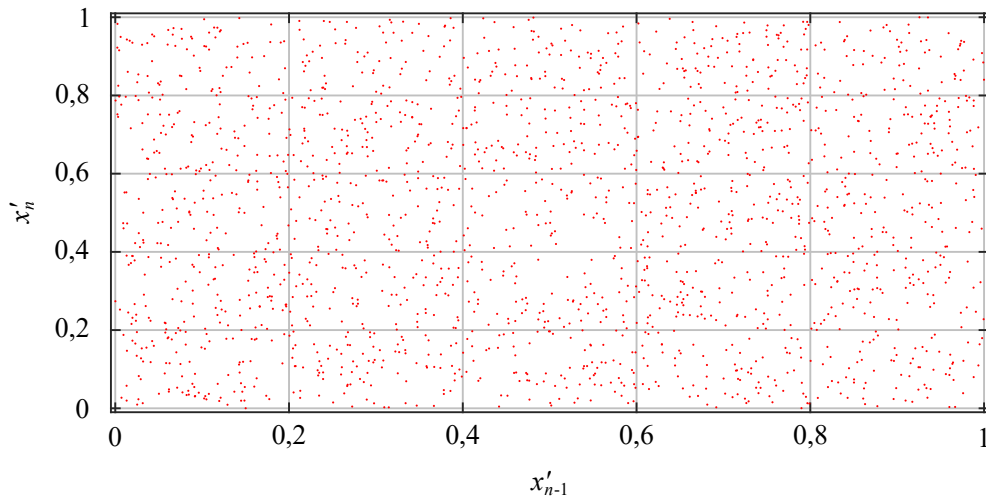
Závislosť medzi hodnotami aktuálneho iterátu  $x_n$  a nasledujúceho iterátu  $x_{n+1}$  je pri aplikáciách v oblasti kryptografie nemysliteľná. Z tohto dôvodu bol pre chaotické šifrovacie algoritmy navrhnutý špecifický kvantizačný algoritmus. Pred tým, ako priradí iterátom  $x_n \in (0; 1)$  hodnotu z množiny  $\{0, 1, 2, \dots, 255\}$ , vynechá prvé štyri desatinné miesta iterátu  $x_n$ . Táto jednocestná operácia zaručí, že z aktuálneho iterátu nebude možné získať hodnotu predošlého iterátu [12].

Vynechanie prvých štyroch desatinných miest bolo zvolené ako kompromis medzi rovnomerným rozložením kvantovaných hodnôt v množine  $\{0, 1, 2, \dots, 255\}$  a čo najnižšou výpočtovou náročnosťou. Bližšie sa tejto problematike venujeme v [11], [12]. Matematický predpis tohto kvantizačného algoritmu je uvedený ako (3):

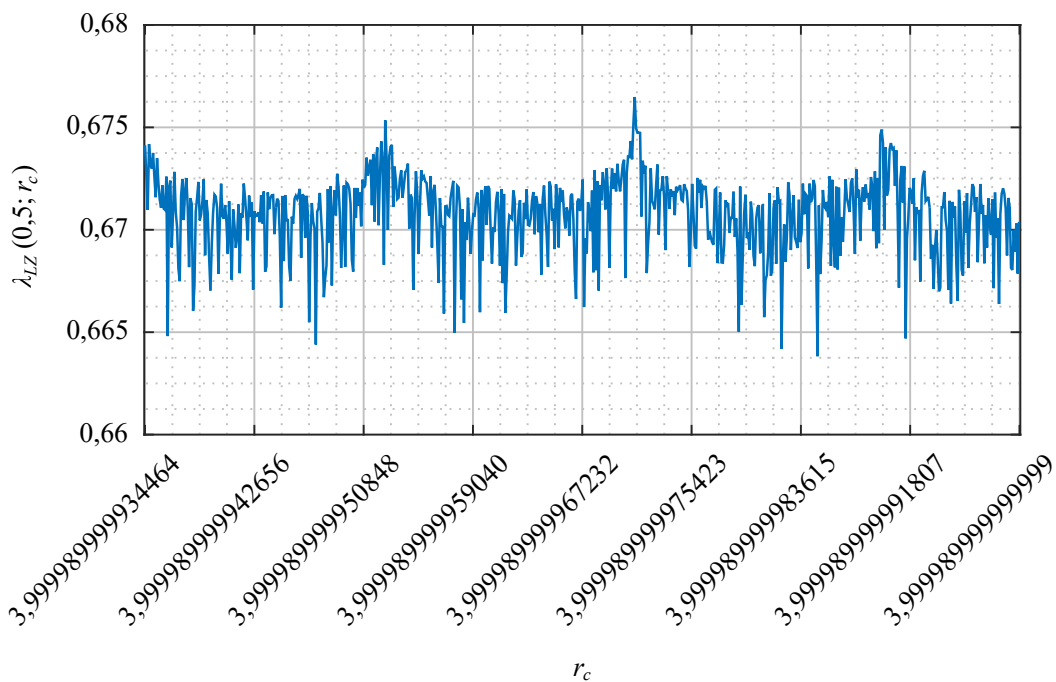
$$x'_n = \lfloor 256 \cdot 10^4 \cdot x_n \pmod{1} \rfloor, \quad (3)$$

kde  $x'_n$  označuje kvantizované hodnoty iterátov,  $x'_n \in \{0, 1, 2, \dots, 255\}$ ,  $x_n$  je postupnosť iterátov generovaných LZ a zátvorky  $\lfloor \cdot \rfloor$  slúžia na zaokrúhlenie na najbližšie menšie celé číslo.

Výsledok použitia predpisu (3) na závislosť susedných iterátov ukazuje Poincarého diagram na Obr. 4. V tomto diagrame boli použité pôvodné hodnoty iterátov pre násobené konštantou  $10^4$  pre umožnenie porovnania s Obr. 2.



Obr. 4 Poincarého diagram pre iteráty LZ po kvantizácii predpisom (3)



Obr. 5 Priebeh hodnôt Ljapunovovho exponentu pre interval hodnôt parametra  $r_c$  (doplňkové hodnoty ku  $r$ )

### C. Obmedzenie vplyvu pevného bodu

Ak sa iterovaním LZ dostaneme k iterátu s hodnotou  $x_n = \frac{1}{1-r}$ , všetky nasledujúce iteráty budú mať rovnakú hodnotu. Táto situácia sa označuje ako tzv. pevný bod zobrazenia. Na riešenie tohto problému sme v minulosti navrhli techniku, ktorá strieda dve rôzne hodnoty parametra LZ. Okrem kľúčom zvolenej hodnoty  $r$  sa používa aj doplnková hodnota  $r_c$ , pričom je dôležité, aby platilo  $\frac{1}{1-r} \neq \frac{1}{1-r_c}$ . Ak táto podmienka platí, výmena  $r$  a  $r_c$  zmení hodnotu pevného bodu a zmenia sa aj hodnoty ďalších iterátov [11].

Pri iterovaní sa hodnota  $r$  používa pre iteráty s nepárnym poradovým číslom ( $x_1, x_3, x_5, \dots$ ) a hodnota  $r_c$  pre iteráty s párnym poradovým číslom ( $x_2, x_4, x_6, \dots$ ). Interval hodnôt pre  $r_c$  bol zvolený tak, aby sa dosiahli čo najvyššie hodnoty  $\lambda_{LZ}$ , ich priebeh neobsahoval výrazné prepady a aby boli hodnoty  $r_c$  ľahko získateľné z hodnôt  $r$ . Nakoniec sa pre túto úlohu použil vzťah (4):

$$r_c = r - 10^{-5}. \quad (4)$$

Pri použití vzorca (4) sa zabezpečí, že pevné body pre hodnoty  $r$  a  $r_c$  sú dostatočne odlišné aj po vynechaní prvých štyroch desatinných miest pri kvantizácii rovnicou (3) [11]. Priebeh hodnôt  $\lambda_{LZ}$  pre  $r_c \in \langle 4 - 10^{-5} - 65536 \cdot 10^{-15}; 4 - 10^{-5} - 10^{-15} \rangle$  ukazuje Obr. 5. Rýchlosť divergencie (odklonu) od pevných bodov LZ ilustrujú hodnoty v Tabuľke 1.

Tabuľka 1  
Ilustrácia divergencie LZ od pevných bodov pri striedaní hodnôt  $r$  a  $r_c$

$r = 4 - 10^{-15}$ , pevný bod $x_n = 0,75$ $r = 4 - 10^{-5} - 10^{-15}$ , pevný bod $x_n = 0,749999374998438$			
iterát	hodnota	hodnota po vynechaní prvých 4 desatinných miest	hodnota po kvantovaní predpisom (3)
$x_0$	0,75	–	–
$x_1$	0,75	0,999999999998181	255
$x_2$	0,749998125000000	0,981250000002547	251
$x_3$	0,750003749985937	0,037499859367927	9
$x_4$	0,749990624990627	0,906249906268386	231
$x_5$	0,750018749667183	0,187496671829649	47
...	...	...	...
$x_{50}$	0,202236261140394	0,362611403942537	92
$x_{51}$	0,645347023281394	0,470232813938310	120
$x_{52}$	0,91549468250523	0,94682505234301	242
$x_{53}$	0,309456675088959	0,566750889588548	145
...	...	...	...

Za zmienku stojí skutočnosť, že pri vynechaní prvých štyroch desatinných miest čísla 0,75 nezískame číslo 0, ale číslo 0,999999999998181. Je to zapríčinené tým, že namiesto čísla 0,75 sa používa jeho najbližšia reprezentácia v dátovom type double precision, ktorá má približnú hodnotu 0,7499999999998181 [11].

#### IV. ZÁVER

V tomto príspevku sme prezentovali niektoré z techník, ktoré sa dajú použiť na zlepšenie vlastností logistického zobrazenia. Aj keď sa v súčasnosti vie o viacerých jeho nedostatkoch, v oblasti obrazových šifrovacích algoritmov je stále populárne. Práve preto je potrebné skúmať nové techniky, ktoré potláčajú jeho nevýhody.

V budúcnosti sa chceme zamerať na ešte komplexnejšiu analýzu logistického zobrazenia a jeho porovnanie s viacrozmernými chaotickými systémami. Pri použití kombinácie techník na vylepšenie jednoduchších zobrazení je možné uvažovať o priblížení sa ku vlastnostiam komplikovanejších zobrazení pri výrazne nižšej výpočtovej náročnosti.

#### POĎAKOVANIE

Táto práca bola podporená výskumnými projektami VEGA 1/0584/20 “UWB senzorové systémy monitorovania osôb pracujúce v reálnych podmienkach (ReConLoc)” a FEI-2021-71 “Metódy zabezpečenia dát v moderných komunikačných systémoch”.

#### LITERATÚRA

- [1] R. May, Simple Mathematical Models with Very Complicated Dynamics, *Nature*, 1976, roč. 261, č. 5560, s. 459–467. ISSN: 0028-0836. DOI: 10.1038/261459a0.
- [2] S. C. Phatak, S. S. Rao, Logistic Map: A Possible Random-Number Generator, *Physical Review E*, 1995, roč. 51, č. 4, s. 3670–3678. ISSN: 2470-0045. DOI: 10.1103/PhysRevE.51.3670.
- [3] J. Gleick, *Chaos: Making a New Science*. Londýn: Vintage Books, 1998, 380 s. ISBN: 978-07-4938-606-1.
- [4] D. Arroyo, G. Alvarez, V. Fernandez, On the Inadequacy of the Logistic Map for Cryptographic Applications, v *Proc. of 10<sup>th</sup> Spanish Meeting on Cryptology and Information Security*. Salamanca (Španielsko), 2008, s. 1–6.
- [5] K. J. Persohn, R. J. Povinelli, Analyzing Logistic Map Pseudorandom Number Generators for Periodicity Induced by Finite Precision Floating-point Representation, *Chaos, Solitons & Fractals*, 2012, roč. 45, č. 3, s. 238–245. ISSN: 0960-0779. DOI: 10.1016/j.chaos.2011.12.006.
- [6] F. Özkaynak, A Novel Method to Improve the Performance of Chaos Based Evolutionary Algorithms, *Optik*, 2015, roč. 126, č. 14, s. 5434–5438. ISSN: 0030-4026. DOI: 10.1016/j.ijleo.2015.09.098.
- [7] L. Liu S. Miao, A New Image Encryption Algorithm Based on Logistic Chaotic Map with Varying Parameter, *SpringerPlus*, roč. 5, č. 1, s. 289–300. ISSN: 2193-1801. DOI: 10.1186/s40064-016-1959-1.
- [8] M. Alawida A. Samsudin, J. S. Teh et al., A New Hybrid Digital Chaotic System with Applications in Image Encryption, *Signal Processing*, 2019, č. 160, s. 45–58. ISSN: 0165-1684. DOI: 10.1016/j.sigpro.2019.02.016.
- [9] E. Solak, C. Çokal, O. T. Yıldız et al., Cryptanalysis of Fridrich’s Chaotic Image Encryption, *International Journal of Bifurcation and Chaos*, 2010, roč. 20, č. 5, s. 1405–1413. ISSN: 0218-1274. DOI: 10.1142/S0218127410026563.
- [10] M. Preishuber, T. Hütter, S. Katzenbeisser et al., Depreciating Motivation and Empirical Security Analysis of Chaos-based Image and Video Encryption, *IEEE Transactions on Information Forensics and Security*, 2018, roč. 13, č. 9, s. 2137–2150. ISSN: 1556-6013. DOI: 10.1109/TIFS.2018.2812080.
- [11] J. Oravec, L. Ovseník, J. Turán et al., Mitigating Drawbacks of Logistic Map for Image Encryption Algorithms, *Computing and Informatics*, 2020, roč. 39, č. 6, s. 1249–1280. ISSN: 1335-9150. DOI: 10.31577/cai2020\_6\_1250.
- [12] J. Oravec, J. Turán, L. Ovseník et al., Chaotic Image Encryption Algorithm Robust against the Phase Space Reconstruction Attacks, *Acta Polytechnica Hungarica*, 2019, roč. 16, č. 3, s. 37–57. ISSN: 1785-8860. DOI: 10.12700/aph.16.3.2019.3.3.
- [13] J. Oravec, J. Turán, L. Ovseník, Image Encryption Technique with Key Diffused by Coupled Map Lattice, v *Proc. of 28th Intl. Conf. Radioelektronika*. Praha (Česká republika), 2018, s. 1–6. DOI: 10.1109/RADIOELEK.2018.8376374.
- [14] L. Ovseník, J. Turán, T. Huszanik et al., An Image Encryption Algorithm with Plaintext Related Chaining, *Computing and Informatics*, 2019, roč. 38, č. 3, s. 647–678. ISSN: 1335-9150. DOI: 10.31577/cai\_2019\_3\_647.
- [15] IEEE 754-2019: IEEE Standard for Floating-point Arithmetic. DOI: 10.1109/IEEESTD.2019.8766229.

# Parametrizovaný kvantizačný algoritmus pre obrazové šifrovacie algoritmy

*Jakub ORAVEC, Ľuboš OVSEŇÍK, Norbert ZDRAVECKÝ,  
Maroš LAPČÁK*

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

{jakub.oravec, lubos.ovsenik, norbert.zdravecky, maros.lapcak}@tuke.sk

**Abstrakt** – Tento príspevok sa zameriava na kvantizačný algoritmus vhodný najmä pre dedikované obrazové šifrovacie algoritmy. Jeho hlavnou výhodou je závislosť na parametri, ktorý predstavujú intenzity obrazových prvkov otvoreného obrazu. Použitie rôznych obrazov spôsobuje odlišné hodnoty kvantovaných prvkov. Kombináciou postupnosti takto kvantovaných prvkov a spracovaného obrazu je možné dosiahnuť odolnosť obrazových šifrovacích algoritmov voči diferenčným útokom. Príspevok pomocou jednoduchého príkladu ilustruje aj nelinearitu navrhnutého riešenia, ktorá aj pri malých zmenách vstupných hodnôt môže spôsobiť veľké rozdiely medzi výstupnými hodnotami.

**Príučové slová** – kvantizácia, logistické zobrazenie, obrazové šifrovacie algoritmy

## I. ÚVOD

Pri šifrovaní obrazových dát je možné použiť viacero prístupov. Ak sa rozhodneme pre konvenčné šifrovacie algoritmy, ku ktorým patrí napríklad AES (Advanced Encryption Standard) [1], je potrebné dôkladne poznať vlastnosti zvoleného algoritmu a správne ho implementovať. Pre niektoré aplikácie je vhodnejšie použiť dedikované obrazové šifrovacie algoritmy, ktoré boli navrhované so zreteľom na ich plánované použitie a odolnosť voči konkrétnym útokom. Ku týmto aplikáciám patrí napríklad steganografia [2], [3], zabezpečenie medicínskych obrazov [4], prípadne biometrické systémy pracujúce s obrazovými príznakmi [5].

V oblasti dedikovaných obrazových šifrovacích algoritmov existuje niekoľko špecifických útokov, ku ktorým patrí aj Solakov útok využívajúci koreláciu a redundanciu intenzít obrazových prvkov [6]. Tento útok je možné klasifikovať ako diferenčný útok, keďže sa použitím vstupných obrazov s malými rozdielmi snaží zistiť, aký bude vplyv rozdielov na výsledný zašifrovaný obraz. Kvôli tomuto útoku sa pozmenila dovedy bežná architektúra obrazových šifrovacích algoritmov a začali sa využívať techniky závislé na otvorenom obraze (angl. plain-text related techniques). Neskôr sa objavilo niekoľko publikácií týkajúcich sa kryptoanalýzy obrazových šifrovacích algoritmov [7], ktoré ovplyvnili návrh súčasných obrazových šifrovacích algoritmov.

Od roku 2012 bolo prezentovaných viacero techník závislých na otvorenom obraze [8]–[11]. Niekoľko riešení bolo publikovaných aj v rámci našej výskumnej činnosti [12]–[14]. V tomto príspevku sa budeme bližšie zaoberať návrhom nového spôsobu ako docieľiť závislosť obrazového šifrovacieho algoritmu na otvorenom obraze, a to prostredníctvom parametrizovaného kvantizačného algoritmu.

## II. LOGISTICKÉ ZOBRAZENIE

Logistické zobrazenie (LZ) je v obrazových šifrovacích algoritmoch často používaným nástrojom na generovanie pseudo-náhodných (PN) postupností [15]–[17]. Jednou z jeho nevýhod je skutočnosť, že vygenerované prvky – iteráty patria do intervalu  $(0; 1)$ . Pre jednoduchšie použitie iterátov LZ v obrazových šifrovacích algoritmoch, ktoré pracujú prevažne s intenzitami obrazových prvkov z množiny  $\{0, 1, 2, \dots, 255\}$  je preto vhodné vygenerované iteráty LZ kvantizovať na prvky z uvedenej množiny.

Vhodným predpisom pre kvantovanie iterátov LZ je možné docieľiť aj rovnomernú distribúciu iterátov, prípadne úplne potlačiť závislosť susedných iterátov [13], [14], [18]. V tomto príspevku bude prezentovaný spôsob na kvantizáciu iterátov LZ podľa intenzít obrazových prvkov otvoreného obrazu. Takýto postup zabráni použitiu diferenčných útokov s viacerými otvorenými obrazmi vrátane Solakovho útoku [7], [14].

Matematický predpis LZ je daný ako (1):

$$x_{n+1} = r \cdot x_n(1 - x_n), \quad (1)$$

kde  $x_{n+1} \in (0; 1)$  je hodnota nasledujúceho iterátu,  $r \in (0; 4)$  je hodnota parametra LZ a  $x_n \in (0; 1)$  je hodnota aktuálneho iterátu.

Na kvantizáciu iterátov LZ je vo všeobecnosti možné použiť rovnicu (2):

$$x'_n = \lfloor (M + 1) \cdot 10^S \cdot x_n \pmod{1} \rfloor, \quad (2)$$

kde  $x'_n \in \{0, 1, 2, \dots, M\}$  označuje kvantizované hodnoty iterátov,  $M$  je najvyššia možná hodnota kvantovaného iterátu,  $S$  označuje počet vynechaných desatinných miest pri kvantizácii,  $x_n \in (0; 1)$  je postupnosť iterátov generovaných LZ a zátvorky  $\lfloor a \rfloor$  zaokrúhľia číslo  $a$  na najbližšie menšie celé číslo. Najčastejšie používanou hodnotou  $S$  sú 4 desatinné miesta [12], [14].

### III. NÁVRH RIEŠENIA

Navrhnuté riešenie pozostáva z viacerých krokov:

- kvantovanie iterátov LZ 16 bitmi na prvky z množiny  $\{0, 1, 2, \dots, 2^{16} - 1\}$ ,
- rozdelenie každého kvantovaného iterátu na dve binárne slová (prvé s 8 viac významnými bitmi a druhé s 8 menej významnými bitmi),
- kombinácia binárnych slov s intenzitami obrazových prvkov z otvoreného obrazu,
- spojenie dvojíc skombinovaných binárnych slov do výslednej hodnoty iterátu po kvantizácii.

#### A. Kvantovanie iterátov

Na rozdiel od riešení, ktoré kvantujú 8 bitmi (výsledné hodnoty z množiny  $\{0, 1, 2, \dots, 255\}$ ), sa v tomto prípade použije rovnica (2) s parametrami  $M = 2^{16} - 1$  a  $S = 4$ . Výsledné hodnoty je teda možné vyjadriť 16 bitmi, čo sa využije v ďalšom kroku. Hodnota parametra  $S$  je zvolená na základe poznatkov z nášho predošlého výskumu, kde sa zistilo, že táto hodnota spôsobí rovnomernejšiu distribúciu hodnôt kvantovaných iterátov pri zachovaní ďalších vhodných vlastností postupnosti vygenerovaných iterátov [12], [14].

#### B. Rozdelenie na binárne slová

Každý z kvantovaných iterátov v postupnosti  $x'_n$  sa v tomto kroku rozdelí na dve binárne slová s dĺžkou 8 bitov. Postupnosť binárnych slov, ktoré obsahujú osem viac významných bitov z  $x'_n$  sa označí ako  $x'_{M,n}$ . Osem menej významných bitov z  $x'_n$  sa uloží do postupnosti binárnych slov  $x'_{L,n}$ . Pre tento účel sa použije sústava rovníc (3):

$$\begin{aligned} x'_{M,n} &= \left\lfloor \frac{x'_n}{256} \right\rfloor, \\ x'_{L,n} &= x'_n \pmod{256} \end{aligned}, \quad (3)$$

kde výsledkom operácie  $a \pmod{b}$  je zvyšok po celočíselnom delení čísla  $a$  číslom  $b$ .

#### C. Kombinácia binárnych slov

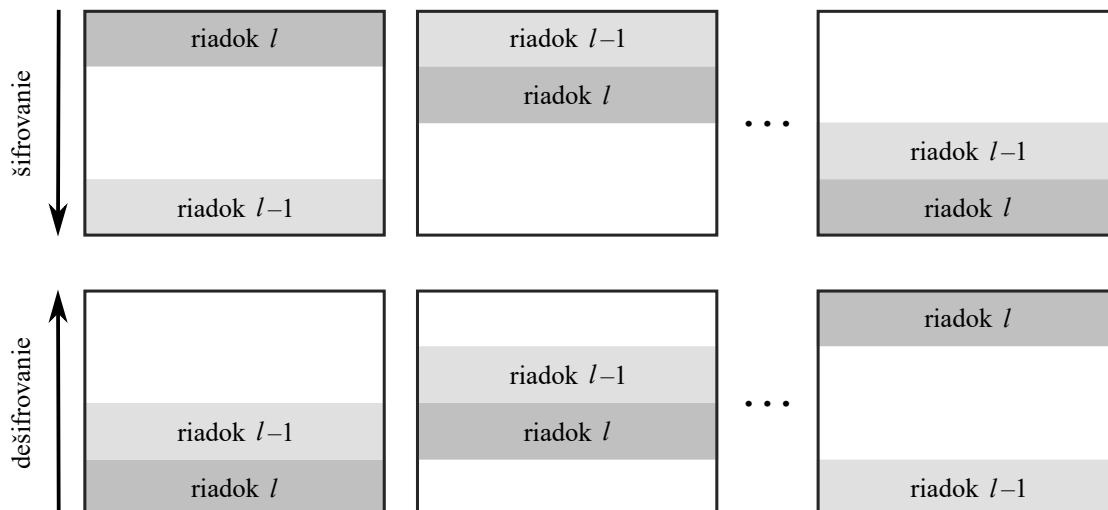
V tomto kroku sa predpokladá, že generovaná PN postupnosť má rovnaký počet prvkov ako matica s intenzitami obrazových prvkov. To umožňuje vykonať kombináciu každej dvojice binárnych slov podľa intenzity jedného obrazového prvku.

Aby bolo možné kombináciu dvojice binárnych slov vykonať pri šifrovaní a zároveň aj pri dešifrovaní, je potrebné zvoliť spôsob, akým sa budú načítavať intenzity obrazových prvkov, ktoré sa použijú ako parameter pri kombinácii dvojíc binárnych slov. Spôsob použitý v navrhnutom riešení je ilustrovaný na Obr. 1. Takáto postupnosť načítavania riadkov intenzít obrazových prvkov zaručí, že rovnaké hodnoty sa získajú pri šifrovaní aj pri dešifrovaní.

Matematický predpis použitý na kombináciu dvojíc binárnych slov a riadkov intenzít obrazových prvkov otvoreného obrazu je možné vyjadriť sústavou rovníc (4):

$$\begin{aligned} Aux_1(l, :) &= P(l - 1, :) \wedge x'_{M,n}(l, :) \\ Aux_2(l, :) &= (255 - P(l - 1, :)) \wedge x'_{L,n}(l, :), \end{aligned} \quad (4)$$

kde matice  $Aux_1$  a  $Aux_2$  sú pomocné matice (z angl. auxiliary), matica  $P$  označuje otvorený obraz (z angl. plain-text image),  $\wedge$  je operátor bitového súčtu,  $l$  je súradnica výšky a dvojbodka : označuje všetky stĺpce v matici  $P$ . Pre prvý riadok zhora ( $l = 1$ ) sa namiesto hodnôt z predošlého riadku použijú hodnoty z posledného riadku. Analogicky sa pri dešifrovaní pre prvý riadok zdola použijú hodnoty z prvého riadku.



Obr. 1 Spôsob načítania intenzít obrazových prvkov pri kvantizácii iterátov LZ

#### D. Spojenie dvojíc binárnych slov

V poslednom kroku sa dvojice binárnych slov v maticiach  $Aux_1$  a  $Aux_2$  zlúčia pomocou súčtu v aritmetike modulo 256 (5):

$$y_n(l, :) = Aux_1(l, :) + Aux_2(l, :) \pmod{256}, \quad (5)$$

kde  $y_n$  je výsledná postupnosť iterátov LZ kvantovaných podľa intenzít obrazových prvkov v otvorenom obraze.

#### IV. EXPERIMENTÁLNE VÝSLEDKY

Vlastnosti navrhnutého riešenia je možné ilustrovať jednoduchým príkladom, v ktorom sa použil obraz v odtieňoch sivej *lena* z databázy USC-SIPI [20]. Z tohto obrazu sa náhodne vybralo päť intenzít za sebou nasledujúcich obrazových prvkov. Tieto boli použité ako parameter pri kvantizácii iterátov LZ, ktoré boli vygenerované s  $r = 4 - 10^{-15}$ . Získané výsledky sú uvedené v Tabuľke 1.

Z Tabuľky 1 vyplýva niekoľko záverov. Prvým je skutočnosť, že hodnoty  $x'_{M,n}$  sú zhodné s hodnotami získanými po kvantizácii podľa (2) s  $M = 255$  a  $S = 4$ . Je to pochopiteľné, keďže obidve postupnosti získavajú ich prvky z ôsmich najvýznamnejších bitov binárnych slov s dĺžkou 8, resp. 16 bitov. Tieto slová majú stále rovnakú osmicu najvýznamnejších bitov.

Pri navrhnutom riešení majú na výsledné hodnoty  $y_n$  vplyv aj zvyšné bity z binárneho slova s dĺžkou 16 bitov. Miera, akou sa prejavia, závisí na intenzitách obrazových prvkov otvoreného obrazu  $P$ , čo je zreteľné aj na hodnotách  $Aux_1$  a  $Aux_2$ . Táto skutočnosť je zjavná najmä v treťom a štvrtom stĺpci tabuľky, kde aj pri podobných hodnotách  $x'_{M,n}$  a  $P$  vzniknú výrazne odlišné hodnoty  $y_n$  (153 a 84).

Tabuľka 1  
 Príklad kvantizácie iterátov so zavedením závislosti na otvorenom obraze

$x_n$	$\sim 0,35107337$	$\sim 0,91128344$	$\sim 0,32338374$	$\sim 0,87522678$	$\sim 0,43681945$
$x'_n$ podľa (2), $M = 255, S = 0$	89	233	82	224	111
$x'_n$ podľa (2), $M = 255, S = 4$	187	213	214	68	49
výsledky navrhnutého riešenia					
$P(l-1, :)$	133	129	124	126	132
$x'_n$ podľa (2), $M = 2^{16} - 1, S = 4$	48 084	54 680	54 876	17 551	12 746
$x'_{M,n}$	187	213	214	68	49
$x'_{L,n}$	212	152	92	143	202
$Aux_1(l, :)$	129	129	84	68	0
$Aux_2(l, :)$	80	24	0	129	74
$y_n$	209	153	84	197	74

Vplyv otvoreného obrazu je zreteľný aj v poslednom stĺpci tabuľky. Pre intenzitu obrazového prvku 132 sa vo výpočte výslednej hodnoty  $y_n$  úplne vynechá 8 najvýznamnejších bitov  $x'_{M,n}$  ( $Aux_1 = 0$ ). Ak by intenzita z  $P$  bola rovná 127, použilo by sa sedem bitov z  $x'_{M,n}$  a iba jeden bit z  $x'_{L,n}$ . V takomto prípade by bola výsledná hodnota  $y_n$  rovná 177. Zmena intenzity v otvorenom obraze o 5 intenzitných jednotiek by teda znamenala zmenu hodnoty prvku kvantovanej postupnosti z 74 na 177.

Navrhnuté riešenie je súčasťou obrazového šifrovacieho algoritmu, ktorý bol predstavený v príspevku [19], kde boli podrobnejšie analyzované ďalšie vlastnosti navrhovaného riešenia. Výsledky uvedené v [19] poukazujú na skutočnosť, že postupnosti generované navrhnutým riešením spĺňajú podmienky dané štatistickými testami NIST 800-22 [21].

## V. ZÁVER

V tomto príspevku sme sa zaoberali parametrizovaným kvantizačným algoritmom, ktorý je možné použiť v obrazových šifrovacích algoritmoch. Princípom tohto riešenia je citlivejšia kvantizácia, rozdelenie získaných hodnôt a ich následná kombinácia podľa intenzít obrazových prvkov z otvoreného obrazu. Výhody riešenia boli predstavené na jednoduchom príklade.

V budúcnosti by sme sa chceli zamerať na ďalšie možnosti vytvorenia závislosti obrazových šifrovacích algoritmov na otvorených obrazoch. Výsledky z našej výskumnej činnosti poukazujú na to, že pri starostlivom návrhu je možné dosiahnuť zaujímavé vlastnosti aj pri pomerne jednoduchých riešeniach.

## POĎAKOVANIE

Táto práca bola podporená výskumnými projektami VEGA 1/0584/20 “UWB senzorové systémy monitorovania osôb pracujúce v reálnych podmienkach (ReConLoc)” a FEI-2021-71 “Metódy zabezpečenia dát v moderných komunikačných systémoch”.

## LITERATÚRA

- [1] FIPS 197: Advanced encryption standard. DOI: 10.6028/NIST.FIPS.197.
- [2] J. Oravec, J. Turán, Substitution Steganography with Security Improved by Chaotic Image Encryption, v *Proc. of Informatics 2017*. Poprad (Slovensko), 2017, s. 284-288. DOI: 10.1109/INFORMATICS.2017.8327261.
- [3] D. Fang, S. Sun, A New Scheme for Image Steganography Based on Hyper-chaotic Map and DNA Sequence, *J. of Information Hiding and Multimedia Signal Processing*, 2018, roč. 9, č. 2, s. 392-399. ISSN: 2073-4212.
- [4] X. Chen, C.-J. Hu, Adaptive Medical Image Encryption Algorithm Based on Multiple Chaotic Mapping, *Saudi J. of Biological Sciences*, 2017, roč. 24, č. 8, s. 1821-1827. ISSN: 1319-562X. DOI: 10.1016/j.sjbs.2017.11.023.
- [5] F. Abundiz-Pérez, C. Cruz-Hernández, M. A. Murillo-Escobar a kol., A Fingerprint Image Encryption Scheme Based on Hyperchaotic Rössler Map, *Mathematical Problems in Engineering*, 2016, roč. 2016, č. 5, s. 1-15. ISSN: 1024-123X. DOI: 10.1155/2016/2670494.
- [6] E. Solak, C. Çokal, O. T. Yildiz et al., Cryptanalysis of Fridrich's Chaotic Image Encryption, *International Journal of Bifurcation and Chaos*, 2010, roč. 20, č. 5, s. 1405-1413. ISSN: 0218-1274. DOI: 10.1142/S0218127410026563.
- [7] M. Preishuber, T. Hütter, S. Katzenbeisser et al., Depreciating Motivation and Empirical Security Analysis of Chaos-based Image and Video Encryption, *IEEE Transactions on Information Forensics and Security*, 2018, roč. 13, č. 9, s. 2137-2150. ISSN: 1556-6013. DOI: 10.1109/TIFS.2018.2812080.
- [8] A. Kanso, M. Ghebleh, A Novel Image Encryption Algorithm Based on a 3D Chaotic Map, *Communications in Nonlinear Science and Numerical Simulation*, 2012, roč. 17, č. 7, s. 2943-2959. ISSN: 1007-5704. DOI: 10.1016/j.cnsns.2011.11.030.
- [9] Y. Zhang, A Chaotic System Based Image Encryption Algorithm Using Plaintext-related Confusion, *Indonesian Journal of Electrical Engineering and Computer Science*, 2014, roč. 12, č. 11, s. 7952-7262. ISSN: 2502-4752.
- [10] M. A. Murillo-Escobar, C. Cruz-Hernández, F. Abundiz-Pérez a kol., A RGB Image Encryption Algorithm Based on Total Plain Image Characteristics and Chaos, *Signal Processing*, 2015, č. 109, s. 119-131. ISSN: 0165-1684. DOI: 10.1016/j.sigpro.2014.10.033.
- [11] Z. Li, A Novel Plaintext-related Image Encryption Scheme Using Hyper-chaotic System, *Nonlinear Dynamics*, 2018, roč. 94, č. 4, s. 1319-1333. ISSN: 0924-090X. DOI: 10.1007/s11071-018-4426-4.
- [12] J. Oravec, J. Turán, L. Ovseník et al., Chaotic Image Encryption Algorithm Robust against the Phase Space Reconstruction Attacks, *Acta Polytechnica Hungarica*, 2019, roč. 16, č. 3, s. 37-57. ISSN: 1785-8860. DOI: 10.12700/aph.16.3.2019.3.3.
- [13] L. Ovseník, J. Turán, T. Huszaník et al., An Image Encryption Algorithm with Plaintext Related Chaining, *Computing and Informatics*, 2019, roč. 38, č. 3, s. 647-678. ISSN: 1335-9150. DOI: 10.31577/cai\_2019\_3\_647.
- [14] J. Oravec, L. Ovseník, J. Turán et al., Mitigating Drawbacks of Logistic Map for Image Encryption Algorithms, *Computing and Informatics*, 2020, roč. 39, č. 6, s. 1249-1280. ISSN: 1335-9150. DOI: 10.31577/cai2020\_6\_1250.
- [15] R. May, Simple Mathematical Models with Very Complicated Dynamics, *Nature*, 1976, roč. 261, č. 5560, s. 459-467. ISSN: 0028-0836. DOI: 10.1038/261459a0.
- [16] S. C. Phatak, S. S. Rao, Logistic Map: A Possible Random-Number Generator, *Physical Review E*, 1995, roč. 51, č. 4, s. 3670-3678. ISSN: 2470-0045. DOI: 10.1103/PhysRevE.51.3670.
- [17] J. Gleick, *Chaos: Making a New Science*. Londýn: Vintage Books, 1998, 380 s. ISBN: 978-07-4938-606-1.
- [18] D. Arroyo, G. Alvarez, V. Fernandez, On the Inadequacy of the Logistic Map for Cryptographic Applications, v *Proc. of 10<sup>th</sup> Spanish Meeting on Cryptology and Information Security*. Salamanca (Španielsko), 2008, s. 1-6.
- [19] J. Oravec, L. Ovseník, J. Papaj, An Image Encryption Algorithm with a Plaintext-related Quantisation Scheme, *IET Image Processing*, 2021, s. 1-17. ISSN: 1751-9659. DOI: 10.1049/ipr2.12174.
- [20] The USC-SIPI image database [online]. [cit. 2021-05-13]. Dostupné na internete: <http://sipi.usc.edu/database/>.
- [21] NIST SP 800-22 rev. 1a: A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications [online]. [cit. 2021-05-13]. Dostupné na internete: <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-22r1a.pdf>.

# Design of a Backup RF Link for Hybrid FSO/RF System

<sup>1</sup>Maroš LAPČÁK, <sup>2</sup>Norbert ZDRAVECKÝ, <sup>3</sup>Luboš OVSENÍK, <sup>4</sup>Jakub ORAVEC

<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>maros.lapcak@tuke.sk, <sup>2</sup>norbert.zdravecky@tuke.sk, <sup>3</sup>lubos.ovsenik@tuke.sk,  
<sup>4</sup>jakub.oravec@tuke.sk

**Abstract** — We are more often encountered in cities and municipalities with introducing optical networks. However, in some cases, expansion of these systems is very demanding and especially non-economic. In such cases, it is switched to the connections made by the radio of frequency systems. However, these RF systems achieve small transmission speeds and low safety versus optical systems. By creating a hybrid FSO/RF system we can provide almost 100% of the transmission quality. It is the combination of these two systems that can eliminate environmental impacts on these communications networks.

**Keywords** — FSO, helical antenna, hybrid FSO/RF system, RF system.

## I. INTRODUCTION

Free Space Optical (FSO) communication is technology using modulated optical beams to ensure wireless data transmission. Communication in this system can take place on short, medium, but also long distances. This means that the use is possible from several meters up to 7 km. Laser is used to form an optical beam, which means that there is a requirement of direct visibility between devices communicating in this way (LOS, Line of Sight). An environment, in which the beam is spread, can cause a reduction in the availability of the system mainly due to non-homogeneity. This non-homogeneity is mainly due to a various composition on the entire length. There is a beam degradation reducing the quality of the optical signal received. This means that the weather change can cause deterioration of the connection and even in some cases the communication failure [1].

### A. Hybrid FSO/RF Systems

Hybrid FSO / RF system is composed of two systems. One is the FSO system and the other is the radio frequency (RF) system. For hybrid systems like this, we distinguish the primary and secondary system. In this case, the primary system is the FSO system as it provides a much higher transmission speed than a classic RF system. For this reason, the RF system represents a secondary system that provides a redundancy of the FSO system in the event of unfavorable conditions for the primary system. Due to redundancy, these systems use connections with different methods of information transmission. This combination was chosen intentionally, since the biggest problem of the FSO system is the mist, with this atmospheric phenomenon does not affect the functioning of RF systems [2].

## II. DESIGN OF A HELICAL ANTENNA SERVING AS A BACKUP RF LINE FOR HYBRID FSO/RF SYSTEM

The helical antenna is a classic example of the antennas whose barrier is created from the wire. As the name suggests the radiator of this antenna is formed by a wire absorbed in the shape of the roll. The power of this emitter is located at one end where a metal plate is also serving as a ground surface. Based on this description, such a basic model is in the principle of a monopole antenna whose power is placed in a metal plate. When describing these antennas, several basic factors affect the emitter of the helical antenna. These factors are spiral length, spiral diameter, number of threads, ground surface dimensions and driver rotation direction. Using these factors, we can edit antenna radiating and adaptation to the working frequency. The direction of rotation spiral determines the polarization of the antenna. If the spiral of the antenna is rotated in the direction of the left-handed roller, we are talking about the left-hand circular polarization, and if the spiral of the antenna is rotated in the direction of the right-hand tapping,



we are talking about a right-hand circular polarization. Since the helical antenna is like the monopole antenna, it is directly dependent on the size of the ground. This problem can be solved by adding another spiral. Type of helical antenna with another spiral is called a bifilar helical antenna. In the case of bifilar helical antenna, the monopole antenna results to a dipole, because the individual spirals are in the phase, therefore occurs at  $180^\circ$  phase shift. This change ensures that the antenna is not dependent on the size of the metal plate, but this metal plate is used only as a reflector to improve radiating characteristics. Also, if you replace the metal plate by plate circuit board, only one driver will be brought to the antenna and the phase shift will be implemented on the platform plate itself. To adapt to the helical antenna, the use of such antenna has several advantages such as high maximum gain and high resistance to recovered waves due to circular polarization. Helical antennas have a wide use. Their adjustments can have omnidirectional radiation, but by magnification of spiral diameter, we can change this antenna from the omnidirectional antenna to the sectoral antenna, i.e., the radiator will be between  $20^\circ$  to  $180^\circ$ . The omnidirectional helical antennas are often used as monopole antennas, when the antenna is designed to work without a metal plate, with a ground surface on the device to which this antenna is attached. Sectoral helical antennas are often placed in the field of helical antennas, which still increases profit and creates a narrow radiant lobe. Such fields can then be used for space communication [3] – [8].

*A. Helical Antenna Design with 2.4 GHz Frequency*

This helical antenna whose dimensions are described in Table 1 and shown on in Fig. 1, is a bifilar, thus an antenna containing two spirals that are in the opposite phase. This means that the printed circuit board placed under spiral serves as a reflector while there is a phase shift. Dimension of the grounding area is in the Table 1. That describes the copper layer dimensions in this printed circuit board. The copper layer serves as a reflector. The spirals are formed by a copper conductor of 1.3 mm thick. The core of the printed circuit boards is formed by FR-04.

Table 1  
 Helical antenna parameters with 2.4 GHz working frequency.

Parameters	Description
Antenna type	Helical antenna
Work frequency	2.4 GHz
Spiral length	144 mm
Spiral diameter	44 mm
Number of threads	10
Direction of threads	Refrigerous threaded
PCB dimensions	75 x 1 mm (diameter x thickness)
Dimensions of the grounding area	75 x 0.035 mm (diameter x thickness)
Used wire	Copper, FR-04
Power type	SMA Connector for Coaxial Driver Connection

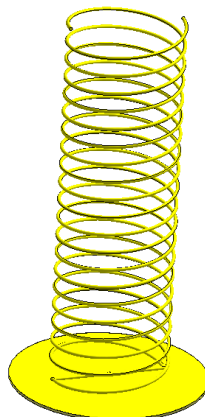


Fig. 1 3D view of the spiral of the helical antenna with 2.4 GHz working frequency.

Based on Fig. 2 we can see good impedance adaptation to this bifilar helical antenna with 2.4 GHz working frequency. This antenna is adapted to a very narrow band because VSWR (Voltage Standing Wave Ratio) values are below 2 in the range of 2.39 GHz to 2.46 GHz, which is precisely unlicensed frequency band in which this antenna will work. At this unlicensed frequency band, the minimum is at a frequency of 2.427 GHz, wherein VSWR reaches 1.38. In the picture it is possible to see additional decreases at frequencies that are resonant for this antenna.

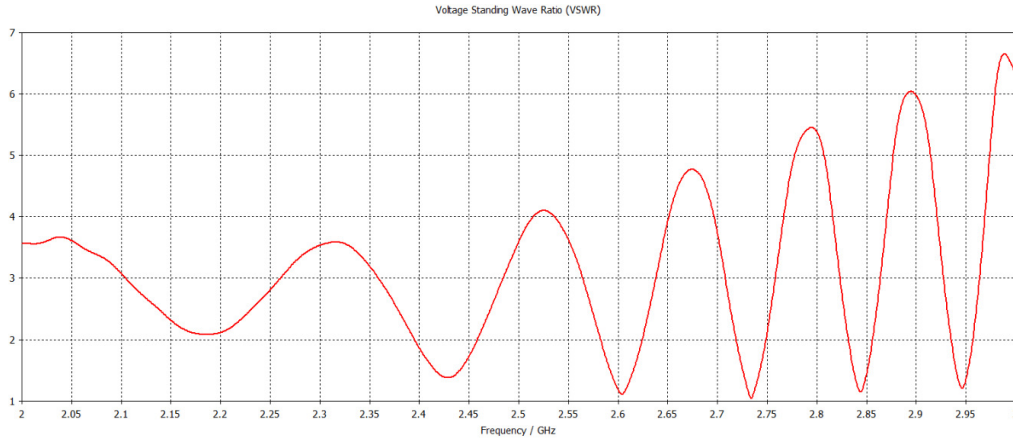


Fig. 2 VSWR helical antenna values with 2.4 GHz work frequency shown in linear chart.

Fig. 3 shows the effectiveness of this helical antenna. It is possible to see its high efficiency whose level is 2.4 GHz to 2.45 GHz more than 90% at a 2.4 GHz working frequency band. The effectiveness of this antenna correlates with what is shown in Fig. 2. Therefore, at the frequencies where VSWR values are located below 2; its efficiency level is high.

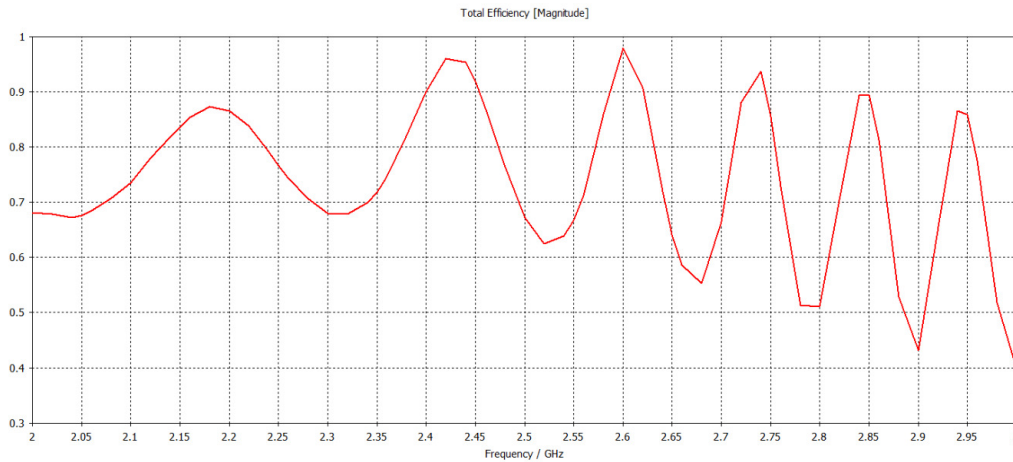


Fig. 3 Efficiency of helical antenna with 2.4 GHz working frequency.

The radiation is shown in Fig. 4 and Fig. 5 at cartesian system  $\varphi = 0^\circ$  and  $\varphi = 90^\circ$  at 2.4 GHz frequency. Fig. 4 shows the radiation characteristics when carbon section  $\varphi = 0^\circ$ . When cutting, the maximum gain is reached 11.7 dBi. In this case, it is also possible to see the direction of maximum profit shown by a dark blue line. The radiation angle width is  $40.1^\circ$ , shown by light blue. Fig. 5 shows the radiation characteristics of angle  $\varphi = 90^\circ$ . In this section, there is also a direction of radiation with maximum gain and can be seen that it is not shifted, and the maximum profit value is unchanged. The radiator angle is at  $41.3^\circ$ . The shape of this antenna is free from significant lateral lobes. Since the efficiency of the antenna on a working frequency over 90% is not necessary to display the value of the direction, as the difference between the direction and gain is minimal in the previous cases.

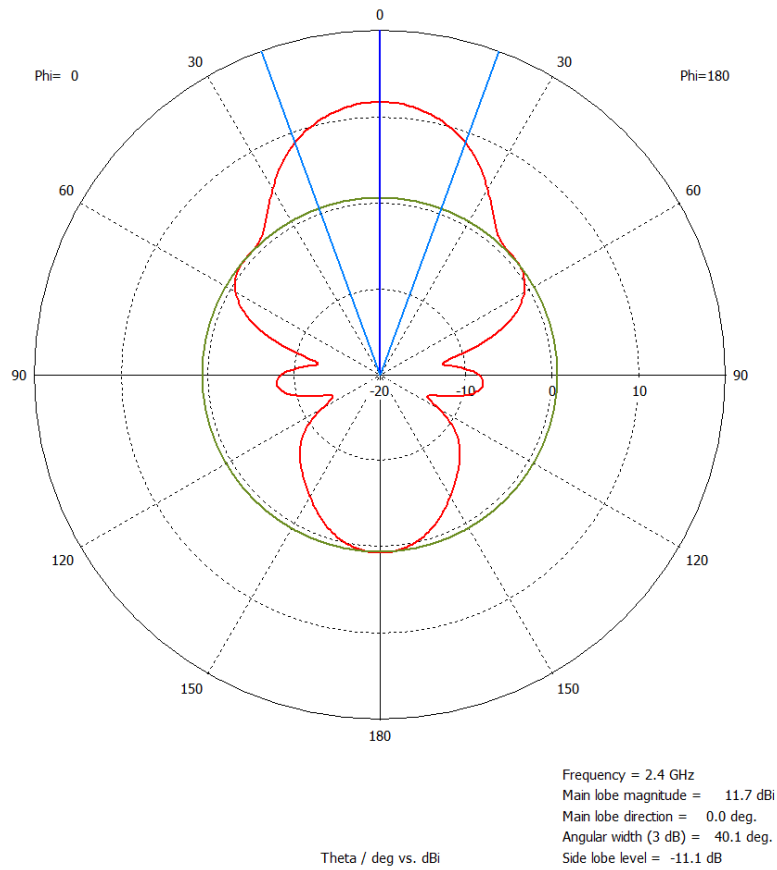


Fig. 4 Radiating characteristics of the helical antenna with a 2.4 GHz working frequency in the polar chart when carbon is cut  $\varphi = 0^\circ$ .

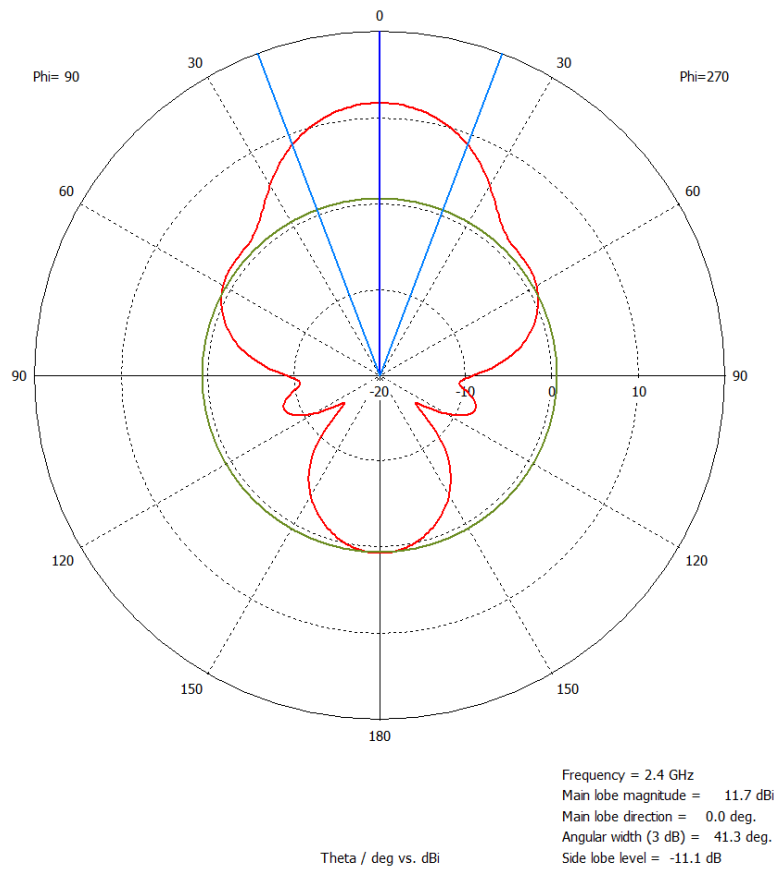


Fig. 5 Radiating characteristics of the helical antenna with a 2.4 GHz working frequency in the polar chart when carbon is cut  $\varphi = 90^\circ$ .

### *B. Results of Simulation*

Based on simulation results, this antenna is highly effective and adequately adapted. Despite the decline in maximum profit and deterioration of radiating angles compared to the funnel antenna, this antenna is a suitable alternative for use, because it has much smaller overall dimensions compared to the funnel antenna, which is one of the advantages of this antenna. Another advantage is the polarization of the antenna since the circular polarization antennas are highly resistant to reflected waves.

### III. CONCLUSION

The aim of the research is to create a proposal for antennas applicable in combination with FSO system and thus create a hybrid FSO/RF system located on the PK13 and L9A building in the TUKE campus. The result is a design and analysis of the helical antenna with a 2.4 GHz working frequency. Its advantages are high level efficiency, high profit level and narrow width of radiation angle. All these benefits are required for hybrid FSO/RF system. According to the results described in chapter II. B we can say that this antenna is ready for deployment into a hybrid FSO/RF system.

### ACKNOWLEDGMENT

This work was supported by research projects VEGA 1/0584/20 “UWB sensor systems for monitoring people working in real conditions (ReConLoc)” and FEEL-2021-71 “Methods of data security in modern communication systems”.

### REFERENCES

- [1] H. A. Fadhil, A. Amphawan, et. al., “Optimization of free space optics parameters: an optimum solution for bad weather conditions,” *Optik - International Journal for Light and Electron Optics*, 2013, vol. 124, no. 19, pp. 3969–3973, DOI: 10.1016/j.ijleo.2012.11.059
- [2] L. B. Stotts, L. C. Andrews, et.al., “Hybrid Optical RF Airborne Communications,” *IEEE*, 2009, vol. 97, no. 6, pp. 1109 – 1127, DOI: 10.1109/jproc.2009.2014969
- [3] A. Z. Elsherbeni, C. J. Reddy, P. Nayeri, “Antenna Analysis and Design using FEKO Electromagnetic Simulation Software,” India: SciTech Publishing, 2014, 245 p., ISBN: 978-1-61353-205-8
- [4] J. R. Hallas, “Basic Antennas: Understanding Practical Antennas and Design,” Connecticut: Amer Radio Relay League, 2009, 183 p. ISBN: 9780872599994
- [5] J. Carr and G. Hippiusley, “Practical Antenna Handbook,” New York: McGraw-Hill/TAB Electronics, 2011, 784 p., ISBN: 9780071639583
- [6] P. Bevelacqua, “List of Antennas,” 2016, Available at: <https://www.antenna-theory.com/antennas/main.php>
- [7] Y. Lo, S. Lee, “Antenna Handbook,” Massachusetts: Morgan Kaufmann Publishers, Van Nostrand Reinhold, 1993, 830 p., ISBN: 0442015933
- [8] W. L. Stutzman and G. A. Thiele, “Antenna Theory and Design,” New Jersey: Wiley, 2012, 843 p., ISBN: 0470576642

# Monitoring Weather Conditions for Hybrid FSO/RF System Using UAV Devices

<sup>1</sup>Maroš LAPČÁK, <sup>2</sup>Norbert ZDRAVECKÝ, <sup>3</sup>Ján TURÁN, <sup>4</sup>Samuel ANDREJČÍK

<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>maros.lapcak@tuke.sk, <sup>2</sup>norbert.zdravecky@tuke.sk, <sup>3</sup>jan.turan@tuke.sk,  
<sup>4</sup>samuel.andrejczik@student.tuke.sk

**Abstract** — Hybrid FSO/RF systems combine FSO and RF systems. Combination of these systems ensures almost 100% functionality. Each of the two systems is affected by different environmental impacts on their functioning. Since the weather can change from an hour to an hour, these environmental impacts need to be recorded. Weather conditions can also be different on the entire transmission route. Consequently, a weather monitoring system is created using UAV (Unmanned Aerial Vehicle).

**Keywords** — FSO, hybrid FSO/RF system, UAV, weather conditions

## I. INTRODUCTION

Optical wireless connections that are also called Free Space Optical (FSO) do not contain many of the restrictions that have microwave connections. FSO is a wireless optical communication technology that uses the effect of optical radiation in free data transmission space between two points without the obstacle to direct visibility between them. Its basic components are transmitters such as lasers or light emitting diodes and signal receivers, such as photodetectors. Fig. 1 shows the FSO communication. FSO is a relatively new technology, but shows many options that allow data to be transferred to up to 8 km with high capacity achieving up to 100 Gb/s. Such systems are two-way thanks to the optical transmitter and receiver at both ends of the connection. Both transmitters and receivers are connected to the data source, such as a computer with a modem. This technology is very sensitive to the type of climate and weather, including physical phenomena. FSO is a direct visibility technology that uses the infrared (IR) light to provide wireless connection between two points. Compared with high-frequency (HF) counterparts, the combination of FSO has a very high optical bandwidth, allowing much higher data speeds in Gigabits per second (GBPS). Because FSO systems use very narrow laser beams, the resulting spatial restriction provides a high reuse factor. In addition, their deployment does not require royalties because they work in the IR optical zone that is not globally regulated [1]. The main disadvantages of these networks include, namely the influence of weather conditions such as rain, fog, or snow in the transmission route [2].

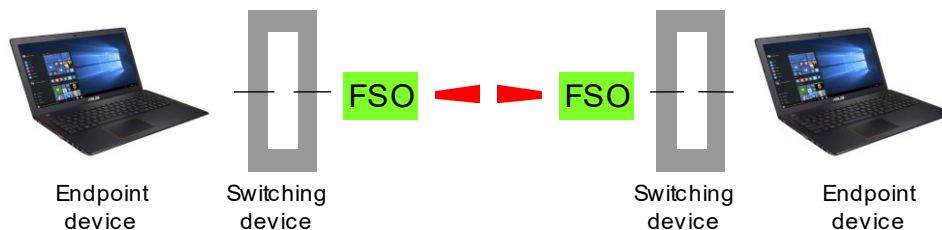


Fig. 1 The principle of operation of the FSO system.

### A. Hybrid FSO/RF System

FSO is one of the least deployed broadband technologies, despite long-term development. It can provide transmission rates comparable to fiber optics without large investment and long

installation. Extending communication deployment, the free environment is mainly limited due to reliability problems. As a result of factors that cause losses, these systems do not reach the required availability of 99.999%, which is called carrier-grade. For this reason, a hybrid FSO/RF communication system as a solution to availability has been designed. It is shown in Fig. 2 [3].

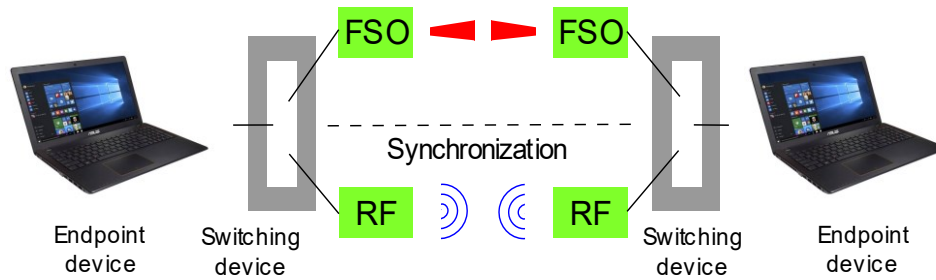


Fig. 2 Principle of operation of a hybrid FSO/RF system.

## II. COLLECTION OF DATA WITH UAV

To obtain data, UAV of a quadcopter type with a vertical take off was used. The specific type of drone was DJI MAVIC as shown in Fig. 3. Drone has 24 powerful computing cores, a transmission system with up to 7 km, 5 optical sensors and 4K camera with a trio-stabilized hinge. It can achieve up to 27 minutes of flight time and maximum distance of 13km thanks to a high performance and efficient engine. When loading the sensor system flight time decreased approximately by a third of the original flight time. Drone provides the possibility to disconnect landing sensors. This ability was used to mount the sensor system to the bottom of the UAV. If this sensor has not turned off, drone would think it is still on the ground and increasing the rotating performance of their engines and could possibly get out of control or fly away from the controller distance [4].



Fig. 3 UVA DJI Mavic Pro.

Due to the short period of time, the UAV had to do several flights. Ten flights at different daily hours were realized for mapping, as large area as possible for testing. Each flight lasted for about 7 minutes. Drone was controlled remotely through the radio frequency driver. Figure from drone tests with the attached sensor system is in Fig. 1.



Fig. 1 UAV device with sensor system.

## III. ANALYSIS OF DATA OBTAINED

We obtained information on air pollution from exploratory flights of a drone in several areas. These data are sufficient for the analysis of air quality and its impact on the FSO communication line. The acquisition of the necessary data has been conducted in the Technical

University of Košice. There were several exploratory flights in this place. The survey was implemented to examine the weather conditions during the transmission route for the FSO system located in the university campus.

It is clear from the measured data that the air pollution level in the TUKE campus is not above average. The temperature during measurement average was 21.3 °C and the air humidity was 40%. The particle concentration values in the air did not exceeded 800 µg/m<sup>3</sup>, which only means the average pollution. The number of measured values in space is well visible in the map in Fig. 2. Individual points displayed on the map represent points of measurements. The measurement was realized every 5 seconds and subsequently allocated to the position based on the time. In Table 1 Google Earth processing data is displayed.

Table 1  
 Sample data suitable for analyzing in graphical interface.

Latitude	Longitude	Temperature (°C)	Humidity (%)	Dust (µg/m <sup>3</sup> )	Icon
49.14767953768352	19.23580693987869	10	30	129	165
49.14807050716106	19.23633649944697	7	37	482	175
49.14859252826276	19.23574298952146	6	35	551	175
49.14783248585005	19.23516762481072	5	38	506	175
49.14781802174114	19.23428932911091	4	39	535	175
49.14877755239755	19.23487037057642	4	43	432	175
49.14735037137988	19.23485361091824	3	44	432	175
49.14891129330181	19.23522872292703	2	48	375	175
49.14691792500138	19.23591501767091	2	49	424	175
49.14642935091275	19.23484989055022	2	52	420	175

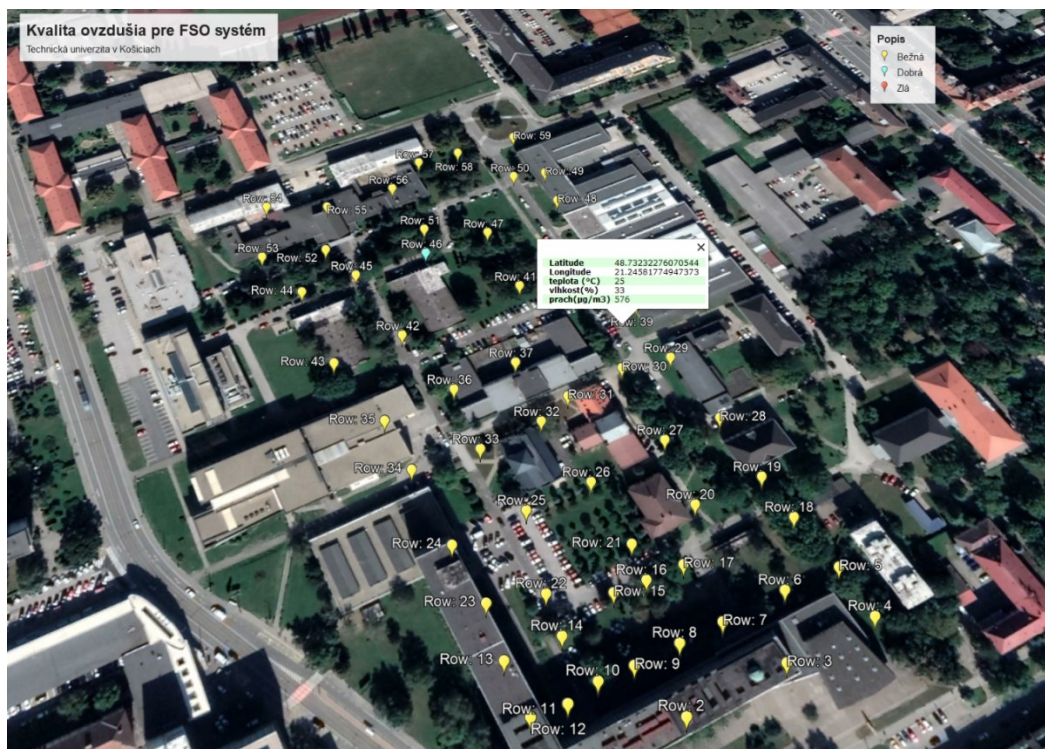


Fig. 2 Display mapped territory in TUKE with recorded values.

Since the influence of air pollution is not too large, it is possible to realize the location of the FSO transmitter and receiver. It is also necessary to consider the layout of buildings and possible obstacles in the beam trajectory. In this case, the first FSO head could be located on the main building of the Technical University (L9) and the second receiver had could be placed on the opposite building of the Mining Faculty (PK19) like that shown in Fig. 3.

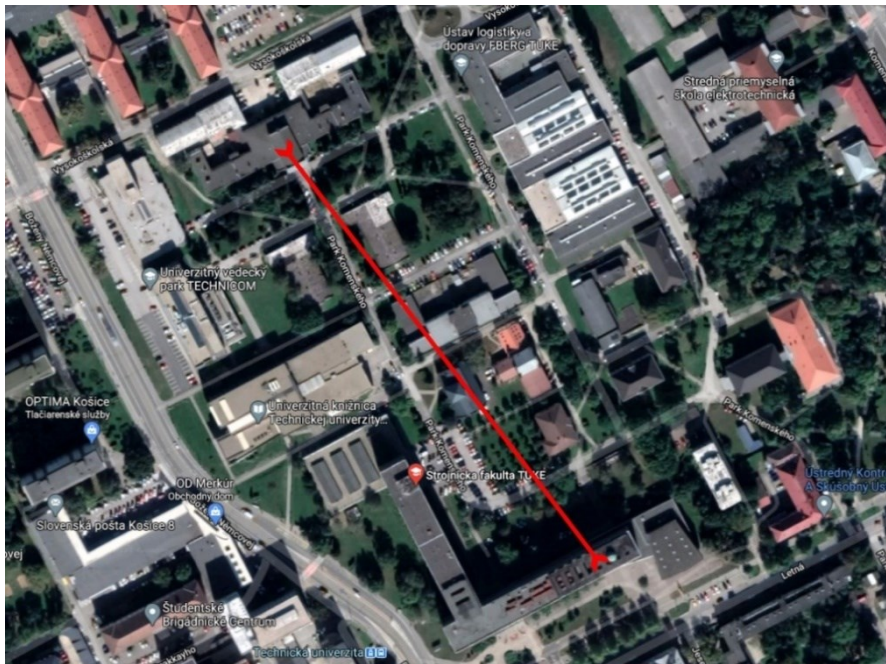


Fig. 3 Design of placement of FSO system in the campus of TUKE.

#### IV. CONCLUSION

In hybrid FSO/RF systems it is necessary to record weather conditions for the hard switching method. If this technology has been used for a longer distance, the effects of weather conditions should also be recorded on the transmission route. Data was obtained from the measurements that were then processed into Google Earth Pro graphical interface. When comparing the data obtained in the graphical interface, it was clear that the University of Technical University in Košice has an average air quality that is suitable for the implementation of FSO lines. In this area, a suitable location of the communication line based on geographic predisposes has been proposed.

#### ACKNOWLEDGMENT

This work was supported by research projects VEGA 1/0584/20 “UWB sensor systems for monitoring people working in real conditions (ReConLoc)” and FEI-2021-71 “Methods of data security in modern communication systems”.

#### REFERENCES

- [1] M. Garlinska, A. Pregowska et. al. “From Mirrors to Free-Space Optical Communication—Historical Aspects in Data Transmission,” in *Future Internet*, 2020, vol. 12, no. 179, pp. 1 – 18. ISSN 1999-5903. DOI: 10.3390/fi12110179.
- [2] J. Singh and N. Kumar, “Performance analysis of different modulation format on free space optical communication system,” 2013, *Optik – International Journal for Light and Electron Optics*, vol. 124, no. 20, pp. 4651 – 4654, DOI: 10.1016/j.ijleo.2013.02.014.
- [3] O. Ugweje, *Radio Frequency and Wireless Communications. Internet Encyclopedia*, 2004. 72 p. DOI: 10.1002/047148296X.tie151.
- [4] P. Drabiňák, “Design of weathering system for a suitable location of FSO lines in the investigated space using UAV,” Diploma thesis, Technical university of Košice, 58 p., 2021.



# Analysis of the Proposed Horn Antenna as a Backup Line for Hybrid FSO/RF System

<sup>1</sup>Maroš LAPČÁK, <sup>2</sup>Norbert ZDRAVECKÝ, <sup>3</sup>Luboš OVSENÍK, <sup>4</sup>Jakub ORAVEC

<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>maros.lapcak@tuke.sk, <sup>2</sup>norbert.zdravecky@tuke.sk, <sup>3</sup>lubos.ovsenik@tuke.sk,  
<sup>4</sup>jakub.oravec@tuke.sk

**Abstract** — This publication deals with the design of a backup RF line for a hybrid FSO/RF system. This system consists of a primary FSO line and a secondary RF line. Both communication lines have different environmental influences on their functionality. In the case of FSO, it is mainly fog and in the case of the RF line, it is rain and mutual interference from users. Therefore, it is necessary to properly design and configure a backup RF line to operate in the unlicensed band and to meet all the conditions for proper functionality.

**Keywords** — FSO, horn antenna, hybrid FSO/RF, RF link

## I. INTRODUCTION

Optical networks can achieve very high transmission speeds. Another advantage is the provision of secure and broadband symmetric transmission. When comparing optical networks with other networks, we can say that the optical network is not dependent on limited and regulated spectrum and frequency licensing, such as radio networks. However, the condition for transmission is direct visibility between the transmitter and receiver. For optical networks, a laser or LED can be used as a transmitter and an optical detector can be used as a receiver. Therefore, optical systems operate in full duplex mode. A disadvantage of FSO (Free Space Optics) networks is the dependence on the effects of the atmospheric transmission environment (smog, fog (especially), snow, etc.) [1].

As mentioned above, FSO junctions mainly suffer from fog attenuation. The RF connection for FSO / RF hybrid systems should therefore meet the following requirements:

- availability during the period of loss of the FSO connection, i.e., during the fog event,
- bandwidth comparable to an FSO link to provide similar performance,
- operation in the frequency band without a license.

As it turns out, these requirements are contradictory. High carrier frequencies must be used to achieve high bandwidths. These frequencies, on the other hand, are strongly affected by precipitation and fog, although the attenuation caused by fog is much lower than in the case of FSO connections. On the other hand, lower frequencies are not affected by fog at all, because unlicensed systems in the ISM bands (2.4 or 5 GHz), are used for IEEE 802.11a/b/g compatible wireless LAN devices, are only affected by rain or multi-user interference [2].

Hybrid FSO/RF link is considered as alternative and cost competitive solution for high-speed, point-to-point wireless communication. They combine the advantages of both links. FSO links provide gigabit data rates and low system complexity, but there are data loss mainly due to fog and scintillation. RF links have lower data rates compared to the FSO links, but their main advantage is their independence from the weather.

## II. DESIGN OF BACKUP RF LINKS FOR HYBRID FSO/RF SYSTEM

The horn antennas belong to the group of aperture antennas and are considered a characteristic example of this group. These antennas consist of two main parts that are waveguide and horn. The shape of the waveguide depends on the shape of a horn. The horn can have either a cone or shape of the needle. This type of antenna also allows various power options or using a waveguide or using a coaxial driver if the antenna is adapted to it. The

positive properties of these antennas include a large bandwidth as suitably adapted horn antennas can have bandwidth of up to 1 GHz. Another advantage is the high directionality of these antennas, but also a high level of efficiency. This means that the horn antennas achieve high profit and due to a small radiation angle, this type of antenna is ranked among directional antennas. Thanks to these features, the horn antennas are used in multiple areas. For example, they are used as standards for calibration, as well as resources for large parabolic reflector antennas or such short-range radar systems [3] – [9].

*A. Proposal of Horn Antenna with 2.4 GHz Working Frequency and its Analysis*

Table 1 shows the dimensions of the horn antenna with a 2.4 GHz employment frequency. The power supply is solved using a SMA connector for connecting the coaxial conductor. A monopole antenna is used as an antenna emitter. The horn skeleton serves as a ground surface and for subsequent radiation antenna. The antenna skeleton is formed of 1mm thick aluminum. The emitter consists of a copper driver with 1mm radius. In Fig. 1 there are several openings and ribs in the horn of this antenna, which further improve the radiation of the antenna but make the manufacturing process more demanding. These holes and ribs are different depending on the side on which they are located.

Table 1 Horn antenna parameters with 2.4 GHz working frequency.

Parameters	Description
Antenna type	Horn antenna
Work frequency	2.4 GHz
Dimensions of waveguide	10.5 x 4.9 x 4.9 cm (width x height x length)
Dimensions of horn	48.33 x 50.4 x 40.4 cm (width x height x length)
Used wire	Aluminum, copper
Power type	SMA Connector for Coaxial Driver Connection

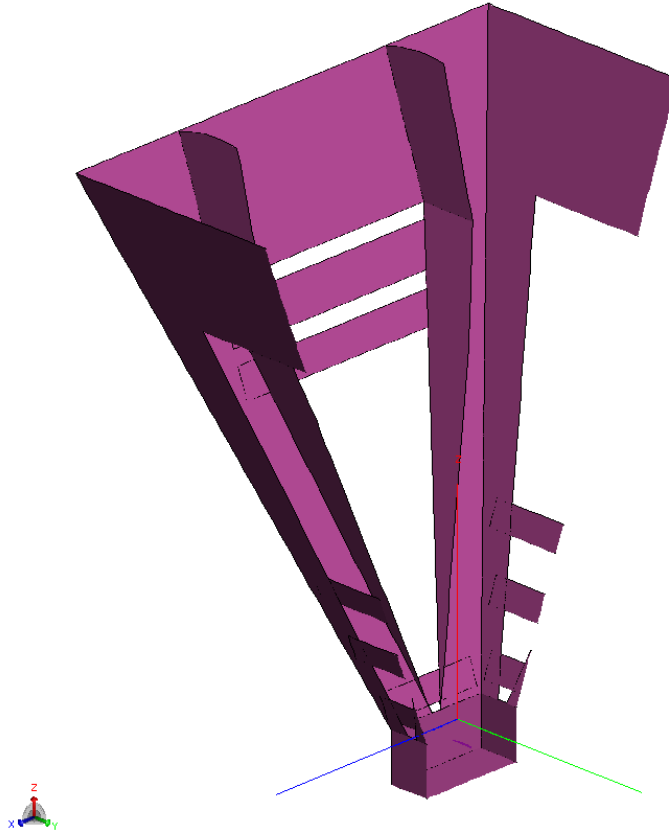


Fig. 1 Section of a horn antenna with a working frequency of 2.4 GHz in 3D display in the FEKO package

From Fig. 1 we can see that the ribs are only found on the sides of the horn that are parallel to the X axis. On these sides of the horn parallel to the X axis we can see four holes that are

trapezoidal shape. The various sides of these holes have the same angle as the horn itself, i.e., their height is enough to describe these holes. The height of the holes on this side is in Table 2. Single holes in this table are sorted from below. In Table 3 are the heights of the openings on the sides of parallel to the axis of Y.

Table 2 The heights of the holes in the horn antenna with a working frequency of 2.4 GHz on the sides parallel to the X axis.

Holes	Height
First hole	1.5 cm
Second hole	27 cm
Third hole	1.5 cm
Fourth hole	0.6 cm

Table 3 The heights of the holes in the horn of the horn antenna with an operating frequency of 2.4 GHz on the sides parallel to the Y axis.

Holes	Height
First hole	3 cm
Second hole	3.5 cm
Third hole	1.5 cm
Fourth hole	22.6 cm

Based on Fig. 2, where the VSWR (Voltage Standing Wave Ratio) values are shown, we can see that this horn antenna is suitably impedance matched for the 2.4 GHz frequency and at the same time for the whole unlicensed frequency band available in Slovakia. The lowest VSWR value is at 2.43 GHz at 1.08. This minimum value is located almost in the middle of the unlicensed frequency band we want to use for this antenna. Based on the appearance of the curve in the chart, it is possible to see a wide range of this horn antenna, as VSWR is located below 2 throughout the frequency range from the frequency of 2.151 GHz to a frequency of 2.667 GHz.

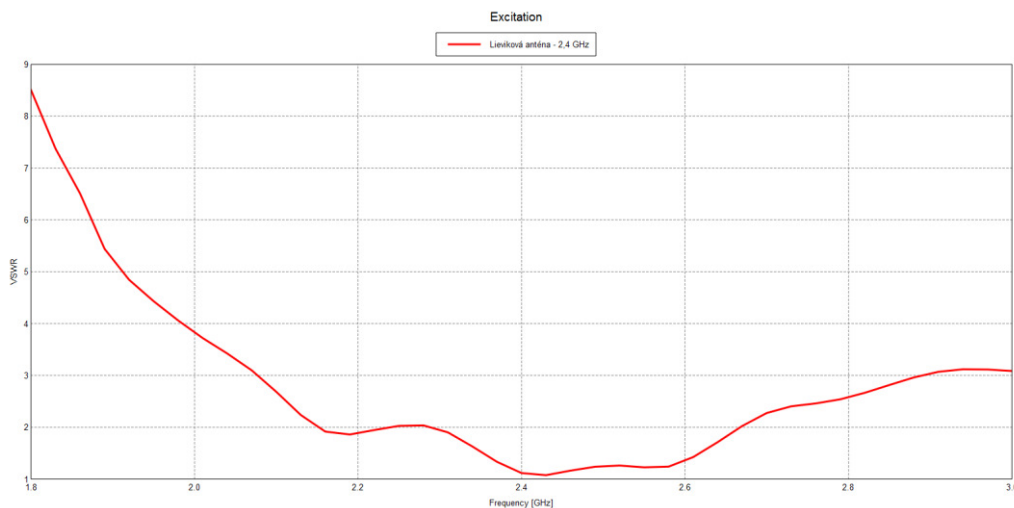


Fig. 2 VSWR values of a horn antenna with an operating frequency of 2.4 GHz in the FEKO package shown in a linear graph.

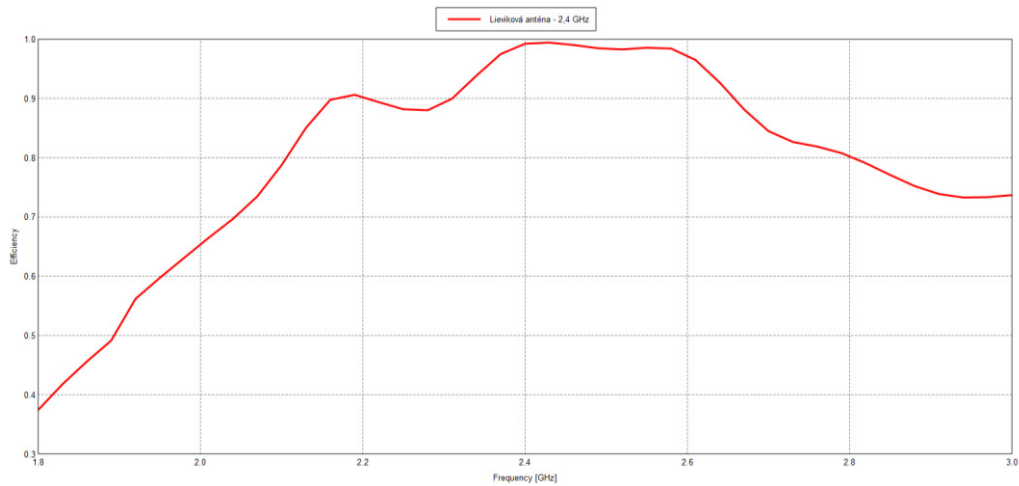


Fig. 3 Efficiency of a horn antenna with an operating frequency of 2.4 GHz in the FEKO package.

The results of this antenna in terms of impedance matching shown in Fig. 2 indicated that the antenna could achieve high efficiency due to the excellent impedance matching of this antenna. As can be seen from Fig. 3, the antenna achieves a high level of efficiency. The maximum efficiency level is 99.4% at a frequency of 2.43 GHz. As mentioned above, the antenna can be considered broadband, as the efficiency level shown in Fig. 3 is at 88% in the same frequency range indicated in this section.

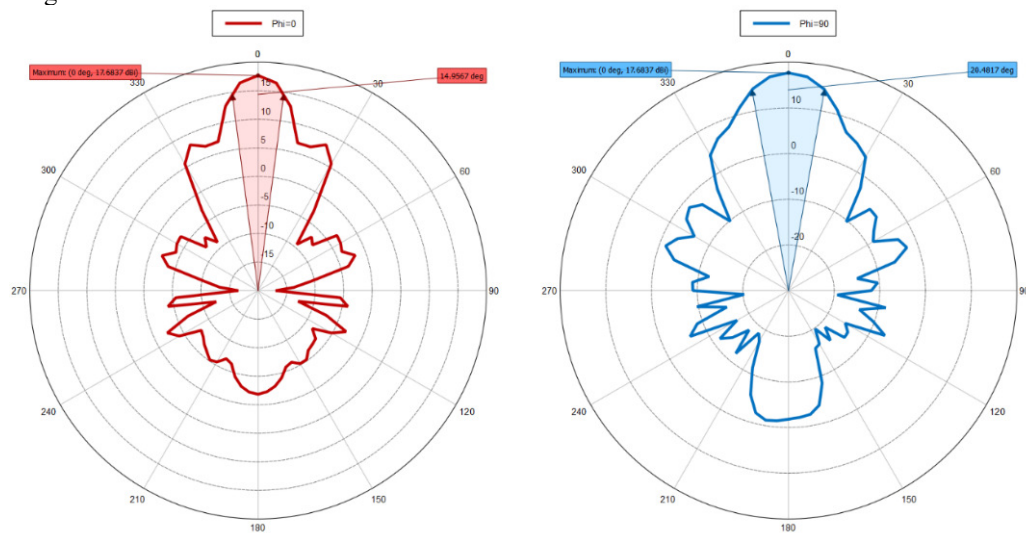


Fig. 4 Radiation characteristics of a horn antenna with an operating frequency of 2.4 GHz in the FEKO package in a polar graph at sections with angles  $\phi = 0^\circ$  and  $\phi = 90^\circ$ .

On the radiating characteristics shown in Fig. 4 at angles  $\phi = 0^\circ$  and  $\phi = 90^\circ$ , we can see the excellent directions of this antenna. The efficiency of this antenna at the working frequency achieves a level of more than 99%, so it is not necessary to indicate the value of the direction, as the maximum profit level of this antenna is almost the same as the level of direction. In the case of angle  $\phi = 0^\circ$ , which is marked with red, the antenna achieves the maximum yield of 17.6837 dBi and the radiator angle is at 14.95°. The second part of Fig. 4, which is displayed by a blue color, shows the radiation characteristics when the carbon cut  $\phi = 90^\circ$ , wherein the antenna reaches a maximum yield of 17.6837 dBi and the radiation angle at 20.48°. The width of the radiation lobe is determined between two points in which the profit is 3 dBi less than in the case of maximum profit. The maximum profit level is the same at both cuts by radiating characteristics. This means that there is no movement of the maximum radiation antenna. In Fig. 4 we can also see the incurred lateral lobes, which do not have a big impact when radiating, as profit in these places is more than 10 dBi lower than maximum profit.

#### B. Evaluation of achieved results

This 2.4 GHz work-frequency antenna reaches good results, as it is adequately adapted for this work frequency. The antenna radiates with high efficiency and thus it is an antenna with a wide range of frequency radiations in which they can effectively radiate. Another advantage of this antenna is its high gain and a small width of the radiant lobe. This antenna also has

disadvantages. The structure is more complicated compared to the classic horn antenna. The biggest problem of this antenna is its dimensions that are very large compared to another antennas.

### III. CONCLUSION

The result is the design and analysis of a horn antenna with an operating frequency of 2.4 GHz. As can be seen from the simulations, this antenna provides excellent parameters such as impedance matching and efficiency and can be used directly as a backup RF line. Its advantage is high efficiency, high gain, and small beam angle. All these benefits are required for a hybrid FSO / RF system. One of the disadvantages of this antenna is its size, which greatly complicates its real use for our system.

### ACKNOWLEDGMENT

This work was supported by research projects VEGA 1/0584/20 “UWB sensor systems for monitoring people working in real conditions (ReConLoc)” and FEI-2021-71 “Methods of data security in modern communication systems”.

### REFERENCES

- [1] R. Pužmanová, “Bezdrátová optika a její možnosti,” 2006, available at: <https://www.dsl.cz/clanky/533-bezdratova-optika-a-jeji-moznosti>.
- [2] O. Ugweje, Radio Frequency and Wireless Communications. *Internet Encyclopedia*, 2004. 72 p. DOI: 10.1002/047148296X.tie151.
- [3] A. Z. Elsherbeni, C. J. Reddy, P. Nayeri, “Antenna Analysis and Design using FEKO Electromagnetic Simulation Software,” India: SciTech Publishing, 2014, 245 p., ISBN: 978-1-61353-205-8.
- [4] J. R. Hallas, “Basic Antennas: Understanding Practical Antennas and Design,” Connecticut: Amer Radio Relay League, 2009, 183 p. ISBN: 9780872599994.
- [5] J. Carr and G. Hippisley, “Practical Antenna Handbook,” New York: McGraw-Hill/TAB Electronics, 2011, 784 p., ISBN: 9780071639583.
- [6] P. Bevelacqua, “List of Antennas,” 2016, Available at: <https://www.antenna-theory.com/antennas/main.php>
- [7] Y. Lo, S. Lee, “Antenna Handbook,” Massachusetts: Morgan Kaufmann Publishers, Van Nostrand Reinhold, 1993, 830 p., ISBN: 0442015933.
- [8] W. L. Stutzman and G. A. Thiele, “Antenna Theory and Design,” New Jersey: Wiley, 2012, 843 p., ISBN: 0470576642.
- [9] T. S. Bird, “Fundamentals of aperture antennas and arrays: from theory to design, fabrication and testing,” New Jersey: Wiley, 2016, 448 p., ISBN: 9781119127451.

# CMS server pre Laboratórium optoelektronických systémov

<sup>1</sup>Samuel ANDREJČÍK, <sup>2</sup>Ján TURÁN, <sup>3</sup>Maroš LAPČÁK

<sup>1,2,3</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>samuel.andrejcek@student.tuke.sk, <sup>2</sup>jan.turan@tuke.sk, <sup>3</sup>maros.lapcak@tuke.sk

**Abstrakt** — Predmetom tohto príspevku je návrh a realizácia CMS servera (WCMS) pre Laboratórium optoelektronických systémov na platforme WordPress LMS servera, ktorý beží na platforme Moodle a tiež virtuálneho dátového úložiska, ktoré je vytvorené na platforme Nextcloud. Primárne sa venuje on-line vzdelávaniu študentov pomocou platformy Moodle.

**Kľúčové slová** — cloud, CMS, LMS, moodle, multimediálny kurz, nextcloud, on-line vzdelávanie, on-line testovanie, wordpress

## I. ÚVOD

Už viac ako rok pretrvávajúca pandémia koronavírusu postavila študentov z triednych stoličiek, vyučujúcim dala miesto kried do rúk počítačovú myš s klávesnicou a postupne ich všetkých premiestnila domov, kde sa od marca 2020 vzdelávajú, resp. on-line vyučujú. Pandémia COVID-19 zastavila vyučovanie na školách zo dňa na deň. Stále nie je jasné, ako dlho budú obmedzujúce opatrenia vlády SR v platnosti. Školy nemajú jednotné inštrukcie ani návod, ako vykonávať online vyučovanie a učители tak stoja pred veľkou výzvou. My sme sa preto rozhodli ísť cestou vytvorenia Content Management System (CMS) servera na platforme WordPress, ako aj Learning Management System (LMS) na platforme Moodle pre online doručovanie obsahu učiva svojimi študentom na KEaMT FEI TUKE.

Pri vypracovávaní tejto práce sme uvažovali o prostriedkoch, ktoré by LOS (Laboratórium optoelektronických systémov) prezentovali na úrovni, nielen pred našimi študentmi, ale aj pred prístupmi z „vonku“. Ako to už býva zvykom, na narastajúci dopyt po kvalitných platformách reagujú aj prostredkovatelia cez svoje ponuky, ktoré túto požiadavku dokážu v súčasnosti dostatočne zasýtiť. Pomocou týchto platforiem je možné vytvoriť pútavé webové stránky, systémy pre on-line vzdelávanie študentov, či dátové úložiská, na ktoré je možné nahráť kapacitne náročnejšie súbory a uchovávať ich pre neskoršie potreby. Z dôvodu zjednodušenia výučby a komunikácie so študentmi sme prejavili záujem o tieto platformy aj my, zamestnanci LOS. V tomto čase trh ponúka široké spektrum riešení na to, aby boli všetky požiadavky splnené.

## II. CONTENT MANAGEMENT SYSTEM – REDAKČNÝ SYSTÉM PRE LABORATÓRIUM OPTOELEKTRONICKÝCH SYSTÉMOV

CMS, u nás častejšie známe ako systémy na manažment obsahu, alebo redakčné systémy slúžia na vytvorenie, správu, resp. optimalizáciu webových stránok. Nespornou výhodou využívania CMS je fakt, že administrátor nutne nepotrebuje ovládať základy programovania a svoju webovú stránku si vie vytvárať za pomoci platformy na to určenej. CMS môžu byť rozšírené prostredníctvom pluginov, rozšírení, tém, šablón, alebo vlastného kódu. Medzi najpoužívanejšie a zároveň najflexibilnejšie CMS dnešných dní z pohľadu využitia prakticky na akomkoľvek serveri určite patria WordPress, Joomla!, Drupal, Wix a ďalšie [1].

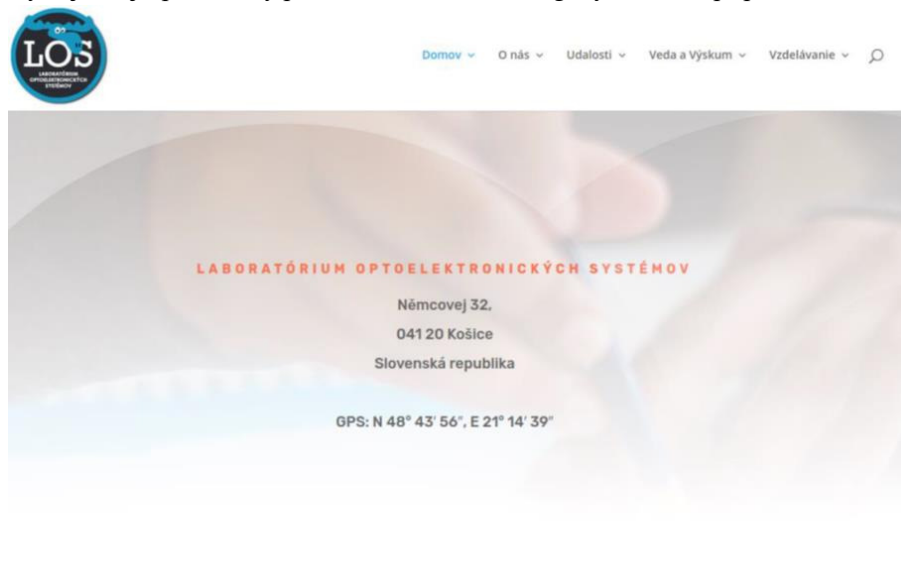
### A. WordPress

Pre potreby výučby a práce so študentmi LOS sme použili redakčný systém WordPress (Obr. 1), ktorý je v súčasnosti najpopulárnejším CMS a beží na viac ako 37% webových stránok celého sveta. V súčasnosti je teda viac ako 450 miliónov webových stránok riešených práve týmto redakčným systémom [2].

WordPress je založený na PHP a pri svojom fungovaní podporuje a využíva databázu

MySQL. Je užívateľsky prívetivý, ľahko sa v ňom dá orientovať a vďaka tak rozsiahlemu pôsobeniu nie je najmenší problém nájsť na internete inštruktážne video, ktoré pomôže v riešení rôznych problémov.

WordPress je veľmi flexibilný a administrátor si ho môže prispôbiť podľa svojich potrieb. V depozitári WordPressu sa nachádza viac ako 56 tisíc pluginov a niekoľko tisíc tém. Takisto poskytuje možnosť si zaplatiť rozšírenia prostredníctvom prémiových doplnkov a tém. Navyše môže byť aj ďalej upravovaný pomocou CSS (Cascading Style Sheet), prípadne časťami kódov.



Obr. 1 Úvodná stránka webu LOS so základnými geografickými údajmi.

### III. LEARNING MANAGEMENT SYSTEM – SYSTÉM NA RIADENIE VÝUČBY PRE LABORATÓRIUM OPTOELEKTRONICKÝCH SYSTÉMOV

Pandémiou koronavírusu sa prakticky roztrhlo vrece so systémami pre on-line vzdelávanie. Pojmy ako dištančné vzdelávanie, či e-learning už nejakú dobu nie sú v slovníkoch študentov vôbec cudzie. Všetky školy postupne prešli na dištančné vzdelávanie svojich študentov, poväčšine však na platformách ako Microsoft Teams, Cisco Webex Meetings, Google Hangout a podobne.

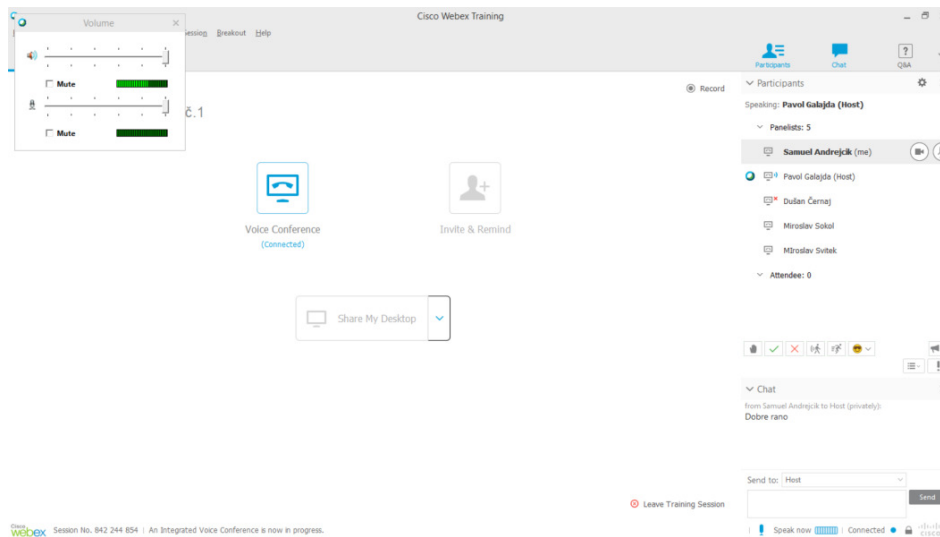
#### A. Cisco Webex Meetings

Na Fakulte elektrotechniky a informatiky Technickej univerzity v Košiciach sa využíva práve aplikácia Cisco Webex Meetings (Obr. 2). Táto platforma umožňuje vzdelávanie študentov dištančne, pričom dovoľuje vyučujúcim v reálnom čase komunikovať so študentmi, prezentovať im prednášané materiály a diskutovať na rozličné témy.

Produkty Cisco Webex poskytujú základné možnosti spolupráce vrátane videohovorov, tímových správ a zdieľania súborov. Táto sada je považovaná za vedúcu platformu spolupráce v zjednotenej komunikácii a je zameraná na spoluprácu v malých skupinách, ale aj na nasadenie v rámci celého podniku, prípadne školy.

Aplikácia je založená na cloudovom princípe webovej aplikácie, poskytuje videokonferenčné služby od spoločnosti Cisco, ktorá umožňuje globálnym virtuálnym tímom spolupracovať na mobilných zariadeniach a štandardných video systémoch v reálnom čase. Webex Meetings obsahuje funkcie ako zdieľanie obrazovky, nahrávanie schôdzky, prispôbitel'né rozloženie a vysielanie schôdzky.

Cisco Webex ponúka vytvorenie samostatnej miestnosti, kde si administrátor môže nastaviť dátum a čas schôdzky a zároveň prizvať všetkých účastníkov. Každá miestnosť funguje na špecifické pozvánky, ktoré prídu účastníkom do ich mailových schránok, resp. je možné nastaviť vstupné heslo, čím sa zamedzí vstupu nežiadaných účastníkov. Počas schôdzky vie administrátor moderovať danú skupinu tým, že dokáže meniť užívateľom ich roly v skupine a tým im obmedziť alebo povoliť určité funkcionality, ktoré aplikácia ponúka. Pri väčšej skupine je vhodné, aby administrátor schôdzky „zamutoval“, teda zablokoval mikrofóny všetkým účastníkom dovedy, kým schôdzku moderuje, aby jeho prednáška nebola rušená okolitými vplyvmi. Rôzne produkty firmy Cisco Webex podporujú rôzny počet účastníkov. Na Technickej univerzite v Košiciach sa využíva sada Cisco Webex Meetings, pre menší počet účastníkov, ale za zmienku stoja aj sady ako Webex Events, alebo Webex Webcasting, ktoré podporujú rozsiahle virtuálne udalosti pre viac ako 3 000, resp. 40 000 účastníkov [3].



Obr. 2 Ukážka využívania aplikácie Cisco Webex v praxi.

### B. Moodle

Väčšina univerzít však okrem týchto platforiem rozbehla na svojich serveroch aj LMS, ktorú poskytuje ďaleko rozsiahlejšie služby, ako len on-line stretávanie vyučujúceho s oštu denťmi. Medzi kľúčové hodnoty, ktoré nesie každý LMS, patrí bezpochyby poskytnutie učebného materiálu študentom aj po skončení vyučovacej hodiny. Študenti sa tak môžu vo vytvorených kurzoch, ktoré predstavujú jednotlivé predmety vrátiť ku prednáškam, resp. výučbovému materiálu, z ktorého majú možnosť opätovne čerpať informácie. Okrem toho sa na LMS systémoch často nachádzajú dokumenty, ktoré musí mať študent so sebou na vyučovaní, väčšinou z praktických cvičení. Prevažne sa jedná o laboratórne protokoly. Vyučujúci tak nemusí pravidelne rozposielať mailom všetky dokumenty, stačí, ak ich uloží na jedno miesto, odkiaľ si ich každý študent má možnosť stiahnuť. Výhodou je fakt, že sa jednotlivé kurzy dajú deliť podľa akademických týždňov, preto sa študent môže ľahko v kurze zorientovať a dostať sa k materiálom, ktoré potrebuje.

Moodle je vzdelávacia platforma navrhnutá tak, aby poskytovala pedagógom, administrátorom a študentom jediný robustný, bezpečný a integrovaný systém na učebné prostredie, ktoré bude upravované podľa požiadaviek.

S aktuálnym počtom viac ako 213 miliónov používateľov je moodle najpoužívanejšou platformou na vzdelávanie. Využívajú ho aj také korporácie ako napr. Shell, či Microsoft, a le využívajú ho aj známe univerzity, napr. Štátna univerzita v New Yorku.

Moodle je navrhnutý tak, aby podporoval vzdelávanie na najvyššej úrovni. Jeho vývoj trvá už viac ako 10 rokov a v súčasnosti poskytuje sadu nástrojov, ktoré sú zamerané na skvalitnenie výučby. Moodle poskytuje funkciu drag-and-drop, disponuje jednoduchým rozhraním a neustále sa vylepšuje, čím uľahčuje výučbu.

Ďalšou nespornou výhodou tejto platformy je fakt, že má pokrytie už vo viac ako 120 jazykoch, vrátane slovenčiny a toto číslo neustále rastie, čím poskytuje pohodlnú orientáciu po platforme ako vyučujúcim, tak aj študentom.

Nakoľko je moodle open-source aplikácia, je možné ho ľubovoľne nakonfigurovať tak, aby spĺňal požiadavky individuálnych potrieb vyučujúcich. Jeho nespornou výhodou je aj fakt, že je neustále podrobovaný aktualizáciám, čím sa stáva veľmi bezpečným systémom. Ochrana osobných údajov a ich bezpečnosť nie je platforme cudzia a jeho návrhári neustále implementujú sady nástrojov, ktoré bránia pred neoprávneným prístupom, stratou, či zneužitím osobných údajov [4].

## IV. CLOUDOVÉ PROSTREDKY PRE IKT VZDELÁVANIE

Vývoj zálohovania úzko súvisí s vývojom sieťových a internetových technológií. Lokálne a globálne siete umožňujú používanie vzdialených počítačov na ukladanie dát po celom svete. Mnoho inštitúcií má tak chránenú svoju infraštruktúru a v prípade akéhokoľvek výpadku, zapríčineného napr. výpadkom energie, vojnou, živelnou katastrofou, a pod., majú svoje dáta zálohované na vzdialených cloudových úložiskách, ktoré sú takto v bezpečí [5].



### A. Nextcloud

Nextcloud by mohol byť prvým krokom nahradeniu služieb, aké ponúka napr. Dropbox, či Skype. V najjednoduchšej podobe by sa server Nextcloud mohol prirovnať k osobnej bezplatnej softvérovej alternatíve k službám ako Dropbox, alebo iCloud. Server je možné nastaviť tak, že budú súbory cez internet vždy k dispozícii, nech už je užívateľ kdekodkoľvek a vďaka platforme je možné sa s nimi podeliť s ďalšími užívateľmi. Okrem toho však Nextcloud ponúka aj ďalšie funkcionality.

Medzi najzákladnejšie dôvody prečo si vybrať Nextcloud bezpochyby patrí fakt, že všetky oficiálne súčasti Nextcloudu sú zadarmo. Jediné, čo sa pri používaní Nextcloudu platí je systémová podpora a aktualizácie. Nextcloud sa v porovnaní s ownCloudom rozhodol sťažiť svoje zameranie na rozširovanie funkcií medzi používateľmi, kým ownCloud je viac zameraný na riešenia vhodné pre firmy a investorov.

Bezpečnosť Nextcloudu sa zdá byť kľúčovým faktorom pri výbere správneho dátového úložiska. Okrem niekoľkých nástrojov, ktoré využíva pre zabezpečenie ochrany údajov používateľov, stojí za zmienku skutočnosť, že pri opakovanej snahe prihlásiť sa nesprávnymi údajmi sa systém pre konkrétne zariadenie zablokuje. Výnimkou je tzv. „brute-force IP whitelist“, teda list IP adries, ktorý dovoľuje istým zariadeniam v liste aj prihlasovanie tzv. „hrubou silou“, resp. mnohokrát opakovanými pokusmi. Z bezpečnostného hľadiska sa však odporúča si takýto list nevytvárať [6].

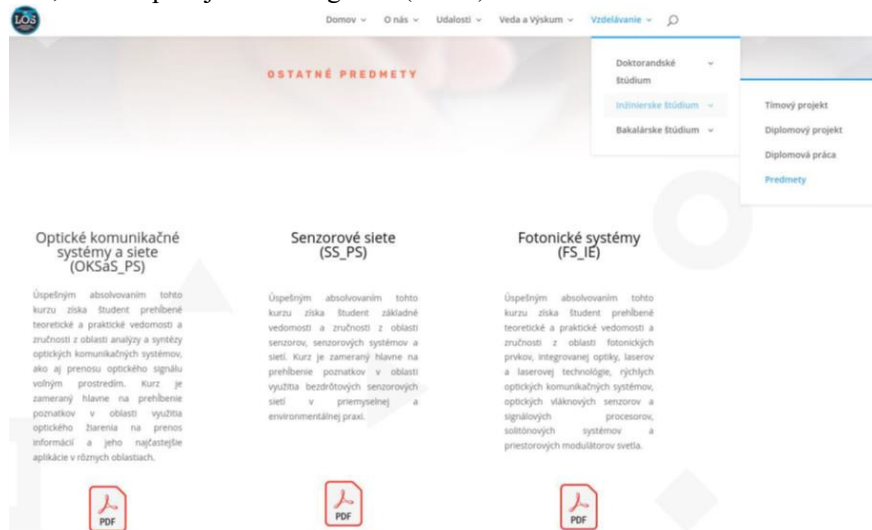
Inštalácia dátového úložiska bola realizovaná na operačnom systéme Linux Ubuntu 18.04. Pred samotnou inštaláciou bolo nutné nainštalovať balík LAMP (Linux, Apache, MySQL, PHP). Následne sa zadaním príkazov do príkazového riadka nainštaloval aj samotný Nextcloud.

Po zadaní prihlasovacích údajov je možné prísť k samotnému dátovému úložisku, nahrávať či sťahovať súbory, ale aj pridávať ďalších užívateľov v sekcii *Settings – Users – New user*.

Aby bolo možné so súbormi v Nextcloude narábať podľa preferencií užívateľa, je nutné ich označiť v poli vľavo od názvu súboru a následne vybrať z poľa *Actions* tú akciu, ktorá je požadovaná.

## V. WEBOVÁ STRÁNKA LABORATÓRIA OPTOELEKTRONICKÝCH SYSTÉMOV

Webová stránka Laboratória optoelektronických systémov mala pôvodne obsahovať ako verejne prístupné, tak aj uzavreté témy a možnosť pre registrovaných používateľov vyjadrovať sa k daným príspevkom. Od tohto plánu sa však upustilo, nakoľko by bolo zbytočne komplikované prihlasovať sa na niekoľko portálov za jediným cieľom – získať potrebné informácie. Z tohto dôvodu je celá webová stránka otvorená každému návštevníkovi a ponúka základné informácie o laboratóriu, ako aj o jednotlivých predmetoch, či zamestnancoch. Práve pri jednotlivých predmetoch je možnosť pokračovať ďalej za účelom získania detailnejších informácií, resp. výukových plánov, či študijných materiálov a to len pre zaregistrovaných užívateľov, ktorí disponujú TUKE loginom (Obr. 3).

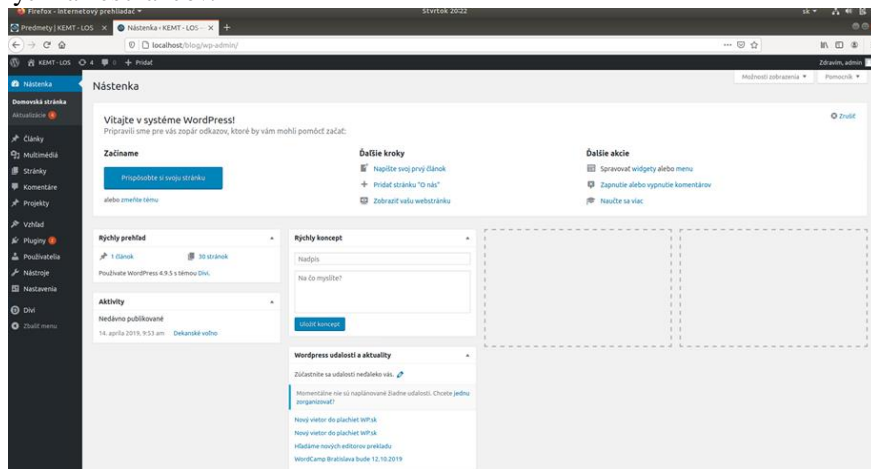


Obr. 3 Karta Predmety predstavujúca predmety LOS vyučované počas inžinierskeho štúdia.

Celý systém je jednoducho prístupný pre administrátora, prípadne viac administrátorov, ktorí môžu kontrolovať a editovať obsah stránky a uverejňovaných informácií, ako aj meniť celkovú

štruktúru a vzhľad stránky (Obr. 4).

Dôvodom vytvorenia webovej stránky je požiadavka, aby všetky informácie o predmetoch vyučovaných zamestnancami Laboratória optoelektronických systémov boli ľahko dostupné z jedného miesta, ako pre potenciálnych uchádzačov o štúdium, tak aj pre študentov, či samotných zamestnancov.



Obr. 4 Administrátorské rozhranie po prihlásení do WordPressu.

## VI. ZÁVER

Základnou úlohou tejto práce bolo zabezpečiť, aby všetky platformy – wordpress, mo o dle a nextcloud boli plne pripravené na spustenie do vyučovacieho systému pre potreby Laboratória optoelektronických systémov. Pri realizácii tejto úlohy sme sa snažili prísť s takým riešením, ktoré bude čo najjednoduchšie pri administrácii inými správcami webovej stránky a zároveň sme sa usilovali o to, aby webová stránka bola čo najprívetivejšia pre oko študenta, ktorý z nej bude čerpať základné informácie a s veľkou pravdepodobnosťou to potenciálnych študentov ovplyvní pri výbere predmetov, ktoré chcú študovať.

Ďalšou úlohou, ktorú sme riešili, bola realizácia výučby na zabehnutom portáli platformy moodle. Na platforme sme postupne pripravili karty pre všetky predmety vyučované na LOS s o všetkými požadovanými kategóriami. Samozrejmosťou je príprava portálu pre finálne skúšky, ktoré je možné vďaka integrovaným testovacím nástrojom. Pre tieto potreby boli vytvorené testovacie otázky a samotné testy.

V rozšírení tejto práce by sme radi pokračovali aj v budúcnosti. Zo všetkých cieľov, na ktorých sa v budúcnosti môže pracovať je napríklad realizácia propagačných videí, ktoré by mali byť zverejnené pri jednotlivých predmetoch, kde by vyučujúci daných predmetov v krátkosti predstavili, čoho sa ich predmet týka a vyzvali tak študentov k zápisu na ich predmety. Veríme, že takáto forma bude mať ešte vyšší dopad na finálny počet študentov, ako je tomu v súčasnosti, len za použitia informačných listov predmetov.

Tento komplexný systém bude možné v budúcnosti rozšíriť a využiť ho tak aj v ďalších laboratóriách, ktoré fungujú, prípadne môžu v budúcnosti vzniknúť na Katedre elektroniky a multimediálnych telekomunikácií.

## POĎAKOVANIE

Táto práca bola podporená výskumnými projektami VEGA 1/0584/20 “UWB senzorové systémy monitorovania osôb pracujúce v reálnych podmienkach (ReConLoc)” a FEI-2021-71 “Metódy zabezpečenia dát v moderných komunikačných systémoch”.

## LITERATÚRA

- [1] What Is a Content Management System (CMS)?. [Online]. 2021. <https://kinsta.com/knowledgebase/content-management-system/>. [Dátum: 14. 05. 2021].
- [2] WordPress vs Other CMS Platforms: How Does WordPress Stack Up Against the Rest? . [Online]. 2018. <https://www.elegantthemes.com/blog/resources/wordpress-vs-other-cms-platforms-how-does-wordpress-stack-up-against-the-rest/>. [Dátum: 14. 05. 2021].
- [3] Cisco Webex [Online]. 2018. <https://searchunifiedcommunications.techtarget.com/definition/Cisco-Webex/>. [Dátum: 14. 05. 2021].
- [4] About Moodle. [Online]. 2020. [https://docs.moodle.org/310/en/About\\_Moodle/](https://docs.moodle.org/310/en/About_Moodle/). [Dátum: 14. 05. 2021].
- [5] B. Trikha, „A Journey from floppy disk to cloud storage“, *International Journal on Computer Science and Engineering* Vol. 02, No. 04 (2010), pp. 1449-1452.
- [6] Nextcloud 13: How to Get Started and Why You Should. [Online]. 2018. <https://www.linuxjournal.com/content/nextcloud-13-how-get-started-and-why-you-should/>. [Dátum: 14. 5. 2021].

# Investigation of amplifiers in DWDM systems

<sup>1</sup> Norbert ZDRAVECKÝ, <sup>2</sup> Luboš OVSENÍK, <sup>3</sup> Maroš LAPČÁK, <sup>4</sup> Jakub ORAVEC

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

<sup>1</sup>norbert.zdravecky@tuke.sk, <sup>2</sup>lubos.ovsenik@tuke.sk, <sup>3</sup>maros.lapcak@tuke.sk, <sup>4</sup>jakub.oravec@tuke.sk

**Abstract** — DWDM (Dense Wavelength Division Multiplex) is a technology that allows the transmission of multiple signals through multiplexing. In this article, we focus on the DWDM system, optical amplifier and 32-channel DWDM system, which was simulated in the OptiSystem program. We evaluate the parameters BER (Bit Error Rate) and Q factor for 40 Gb/s systems with changing amplifiers. We chose EDFA, ROA, SOA amplifiers. The aim is to determine the effect of amplifier changes in DWDM system design. Second goal is to find the optimal amplifier for this design.

**Keywords** — DWDM, EDFA, optical amplifier, ROA

## I. INTRODUCTION

DWDM can be used for a service delivery approach to meet consumer demand for ever-increasing wavelengths. The high demand for increased data rate was recorded due to the increased use of high bandwidth applications. Bandwidth can be increased by installing more optical cables, increasing the bit rate of the system, or multiplexing different wavelengths (WDM). To meet this growing demand for bandwidth, a technique combining different wavelengths, called DWDM, has been developed. Today, this system is used at speeds of 40 and 100 Gb/s. In this work we will compare different designs of optical amplifiers and their influence on transmission limitation in DWDM systems [1][2].

## II. OPTICAL AMPLIFIERS

Optical amplification is very useful in the IM/DD (Intensity modulation/Direct Detection) system to reduce the effect of thermal noise in the circuits and to achieve a high sensitivity of the receiver. In IM/DD modulation scheme where the intensity of an optical source is modulated by the RF or mm-wave signal. Demodulation is achieved through direct detection of the optical carrier and conversion using a photodetector. Erbium-doped fiber (EDFA) devices, which have been developed over two decades, function as elements of concentrated optical gain in terrestrial, submarine and access networks. Today, they support almost all commercial data traffic [3]. Raman amplifiers (ROA) have enabled a significant increase in range and increase in capacity and, unlike erbium-doped devices, are not limited to a specific wavelength band. Raman amplification is based upon stimulated Raman scattering. Semiconductor optical amplifiers (SOA) are designed for signal regeneration and can be classified according to the role they play in optical networks. The terminal amplifier element, referred to as the preamplifier, is used to amplify the weak signal before its conversion from the photonic to the electronic domain. This type provides high gain (> 30 dB), works with a small signal input ( $P_{in} < 25$  dBm) and delivers minimal excessive noise. Repeated optical transmission relies on in-line amplifiers designed to regenerate the signal after each fiber span [4][5]. The in-line amplifier must match the loss induced in the range and the loss of the in-line element, usually providing gains greater than 20 dB. Unlike a single-channel preamplifier, an in-line amplifier is designed and operated as a multichannel amplification block in wavelength and multiplex (WDM) links [6][7]. Fig. 1 shows optical transmission in WDM that is amplified several times.

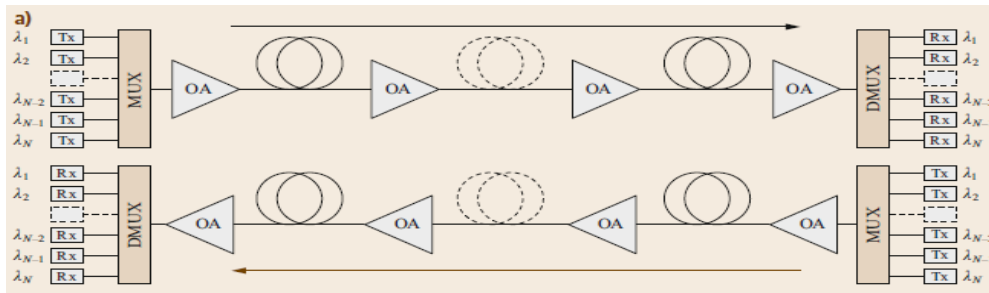


Fig. 1 Repeated amplified transmission of light waves in WDM.

### III. DESIGN OF ULTRA-DWDM SYSTEM WITH OPTICAL AMPLIFIERS EDFA AND ROA

The design consists of lasers, a multiplexer and a demultiplexer, an optical fiber, a dispersion compensating optical fiber (DCF), EDFA amplifiers and optical receivers. The parameters of the 32-channel DWDM of the proposed model are the transmission speed of 40 Gb/s, the central frequency is 193,1 THz, the frequency spacings are 100 GHz. The total optical distance is 240km.

In Fig. 2 is schematic of a DWDM communication system using four EDFA amplifiers with gain 10 dB and four EDFA amplifiers with gain 5 dB.

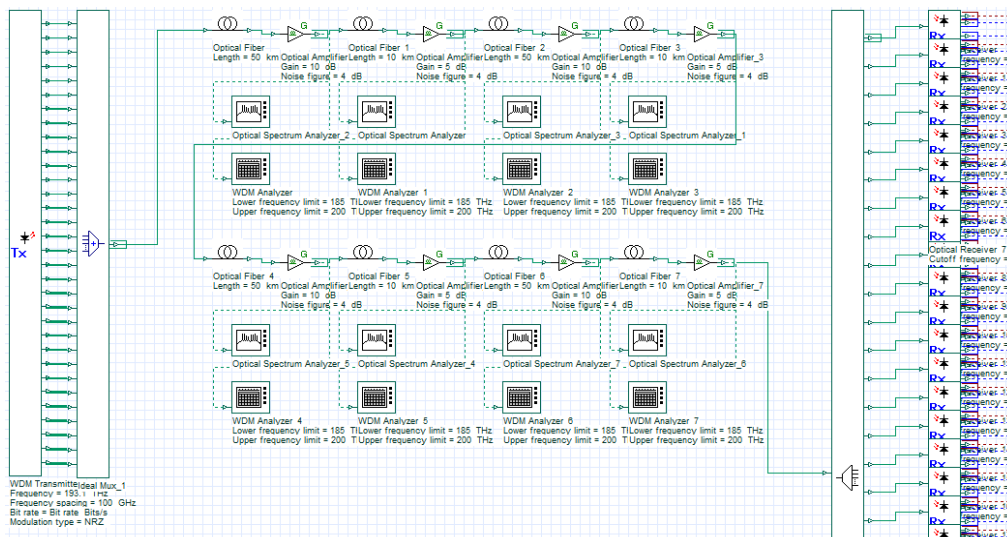


Fig. 2 Schematic of a DWDM communication system using an EDFA amplifiers at a transmission speed of 40 Gb/s

From the eye diagram display, we can extract various parameters of received signal. The most used ones are bit error rate (BER), Q factor measurements measured in decibels and the eye opening. Eye opening signals how much is the analyzed eye diagram opened. The most open part means the best signal-to-noise ratio (SNR). Q factor is a function of SNR and provides a qualitative description of receiver performance. Fig. 3 shows the eye diagrams and (red) BER factor that were simulated for DWDM (EDFA) in the OptiSystem program and Tab. 1 shows simulation results.

Tab. 1 Resulting BER and Q factor values at a transmission rate of 40 Gb/s in DWDM with EDFA (channels 1 and 10)

DWDM channel 1	
Frequency	193,1 THz
Wavelength	1552,5244 nm
BER	$3,22 \times 10^{-63}$
Q factor	16,7374
DWDM channel 10	
Frequency	194,0 THz
Wavelength	1545,3219 nm
BER	$8,86 \times 10^{-32}$
Q factor	11,6693

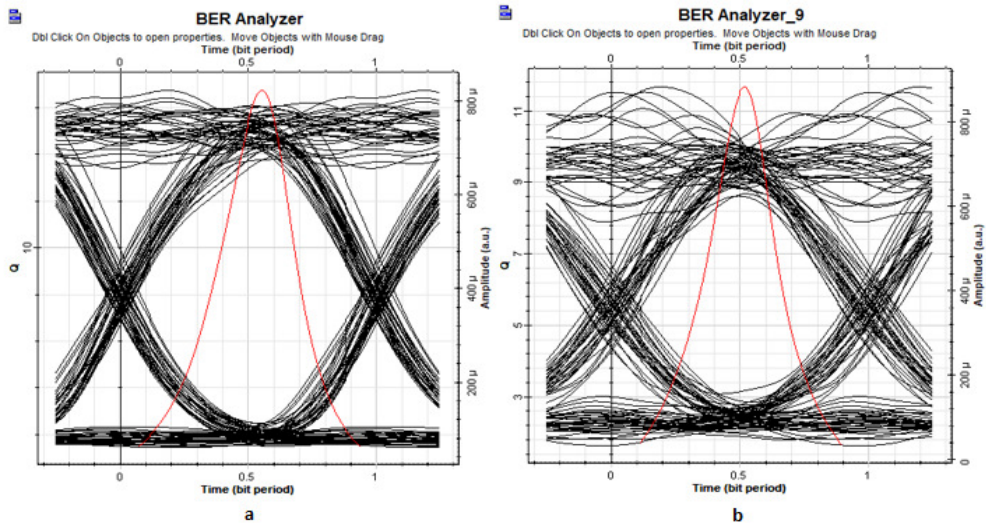


Fig. 3 Eye diagrams at a transmission rate of 40 Gb/s for a.) Channel 1, b.) Channel 13

In Fig. 4 is schematic of a DWDM communication system using ROA amplifier with length 1 km.

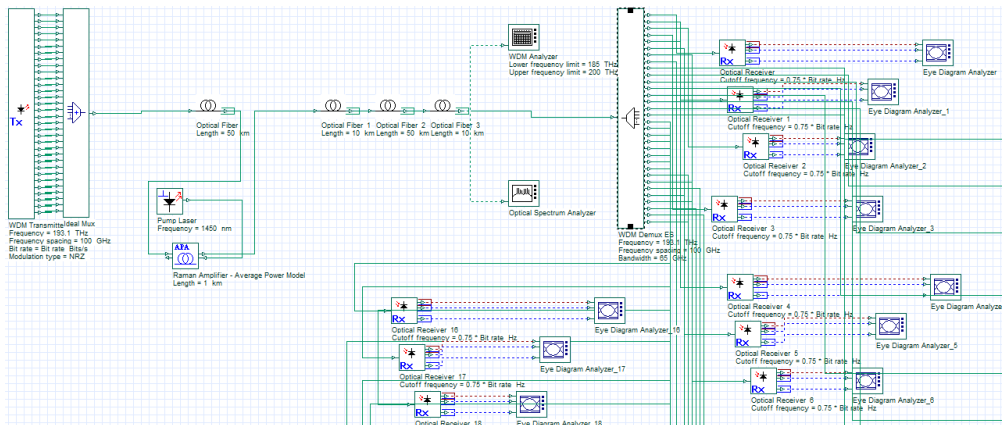


Fig. 4 Diagram of DWDM communication system using ROA amplifier.

In Fig. 5 shows the eye diagrams that were simulated for DWDM (ROA) in the OptiSystem program and Tab. 2 shows simulation results.

Tab. 2 Resulting BER and Q factor values at a transmission rate of 40 Gb/s in DWDM with ROA (channels 1 and 13)

DWDM channel 1	
Frequency	193,1 THz
Wavelength	1552,5244 nm
BER	$2,59 \times 10^{-50}$
Q factor	14,8679
DWDM channel 13	
Frequency	194,0 THz
Wavelength	1545,3219 nm
BER	$5,51 \times 10^{-25}$
Q factor	10,2499

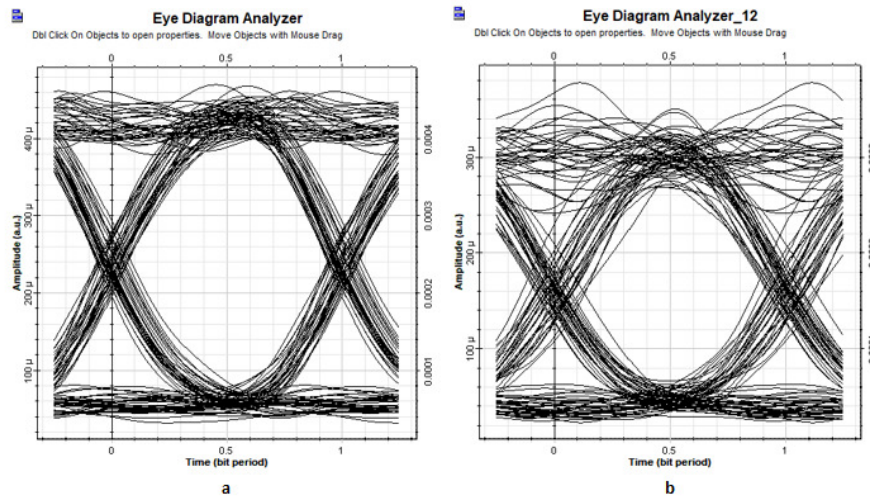


Fig. 5 Eye diagrams at a transmission rate of 40 Gb/s for a.) Channel 1, b.) Channel 13

Our presumption is that the performance of the system with EDFA amplifier will be better, according to the comparison of the results from the simulations. The simulation is performed with components with ideal parameters and without negative effects on the system, such as nonlinear phenomena. The aim of the work was to compare optical amplifiers and not the effects acting on the optical network. In real conditions, the results will not be as good as in the simulation.

#### IV. CONCLUSION

The optical amplifier implements different types of different data. SOA perspective is showing up in recent years, requiring the demand for cheaper, more efficient equipment is the image of our time. SOA performance is still not comparable to EDFA. SOA has higher noise, lower gain, medium polarization dependence and high nonlinearity with fast transition time. The most used is an (EDFA). A DWDM signal is fed into this fiber and a laser source is connected to it at the same time. Optical amplifiers are often used after multiplexing and before demultiplexing. The optical amplifier only amplifies and does not adjust the signal shape (reshape) and time information (retime). By simulating a communication system using an EDFA and ROA amplifier, we found from the measured values that the values of the Q factor and BER are well above the threshold values. In the eye diagram, the intersymbol interference in signal transmission over the length of the optical network was 240 km at a transmission speed of 40 Gb/s. As the baud rate increased, the signal transmission quality deteriorated.

#### ACKNOWLEDGMENT

This work was supported by research projects VEGA 1/0584/20 “UWB sensor systems for monitoring people working in real conditions (ReConLoc)” and FEEI-2021-71 “Methods of data security in modern communication systems”.

#### REFERENCES

- [1] Biswanath Mukherjee, Ioannis Tomkos, Peter Winzer, Yongli Zhao: Handbook of Optical Networks. Switzerland: Gewerbestrasse 11, 6330 Cham, 2020, ISBN: 978-3-030-16250-4, 1169 pp.
- [2] Agrawal, Govind P.: Fiber-Optic Communication Systems (Third Edition). New York: The Institute of Optics University of Rochester, 2002, ISBN: 0-471-22114-7, 546 pp.
- [3] Winzer, Peter J. Energy-Efficient Optical Transport Capacity Scaling Through Spatial Multiplexing. IEEE Photonics Technology Letters. 2011, 23(13), 851 - 853. ISSN 1041-1135. Available on: doi:10.1109/LPT.2011.2140103
- [4] Huszanik Tomáš, Ján Turán and Ľuboš Ovseník. On the Impact of Fiber Nonlinear Effects on the CP-DQPSK Modulated Ultra-DWDM System. Acta Electrotechnica et Informatica. 2019, 19(4), 21-28. ISSN 1335-8243. Available on: doi:10.15546/aei-2019-0026
- [5] Bosco Giovanni, Andrea Carena, Member, IEEE, Curri, Gaudino and Poggiolini, —On the Use of NRZ, RZ, and CSRZ Modulation at 40 Gb/s With Narrow DWDM Channel Spacing, IEEE Journal of Lightwave Technology, Vol. 20, No. 9, Pages 1694-1704, 2002.
- [6] G.694.1: Spectral grids for WDM applications: DWDM frequency grid. ITU-T, 2020. Available on: <https://www.itu.int/rec/T-REC-G.694.1/en>
- [7] Laude, J.Pierre. Dwdm Fundamentals, Components, and Applications: Artech House Optoelectronics Library. London: Artech House, 2002. ISBN 978-1580531771.

# Investigation of modulation in ultra-DWDM systems

<sup>1</sup> Norbert ZDRAVECKÝ, <sup>2</sup> Luboš OVSENÍK, <sup>3</sup> Maroš LAPČÁK, <sup>4</sup> Jakub ORAVEC

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

<sup>1</sup>norbert.zdravecky@tuke.sk, <sup>2</sup>lubos.ovsenik@tuke.sk, <sup>3</sup>maros.lapcak@tuke.sk, <sup>4</sup>jakub.oravec@tuke.sk

**Abstract** — DWDM (Dense Wavelength Division Multiplex) is a technology that allows the transmission of multiple signals through multiplexing. In this article, we focus on the DWDM system, modulations, and 30-channel DWDM system, which was simulated in the OptiSystem program. We evaluate the parameters BER (Bit Error Rate) and Q factor for 40 Gb/s systems with changing modulation. We chose DPSK, DQPSK. The aim is to determine the effect of modulation changes in DWDM system design. Second goal is to find the optimal modulation for this design.

**Keywords** — DPSK, DQPSK, DWDM, optical modulation

## I. INTRODUCTION

The constant development of mobile devices, video streaming, cloud and data centers leads to a significant increase in demands from the customer's point of view. The service provider is therefore looking for ways to resolve these claims. Demands such as increasing the capacity of the channel, increasing the speed of downloading or uploading data, greater network efficiency is needed nowadays. The DWDM system normally operates at a transfer rate of 10 to 40 Gb/s. A coherent 100 Gb/s solution is a great choice with ultra-high bandwidth and cost efficiency. It is used for linear solutions and is suitable for long-distance applications.

## II. OPTICAL MODULATION

One of the basic parameters of optical fiber communication systems (FOC), which largely determine the parameters of a given system, is the method of modulation of the optical signal. The importance of modulation lies in influencing the carrier optical signal by the information-carrying signal so that the resulting modulated signal can transmit the information over the greatest possible distance at the highest possible speed and at the lowest possible cost.

Advanced modulations, which have caused a big leap in modern communications, have come to the fore mainly due to significantly improving utilization and capacity of the channel. Another significant plus is their improved performance in long-distance fiber optic communication. Especially due to the significantly improving properties over long distances, these technologies have started to be used and are often sought after [1][2]. In modern communication, modulation formats of various kinds are used, for example: amplitude shift keying (ASK), frequency shift keying (FSK), phase shift keying (PSK), quadrature phase shift keying (QPSK), differential quadrature phase shift keying (DQPSK). The binary sequence is transmitted using an amplitude key change, referred to as "on-off" keying (OOK) [3][4]. In optical communication, a Mach-Zehnder interferometer is used, for example, as a demodulator of DPSK and DQPSK.

Differential Phase Shift Keying (DPSK) is a fast and stable modulation format and is used as a tool for many optical applications. The main advantage is lower error rate compared to simple modulations and it is not necessary to know the whole absolute phase. DPSK is a type of phase modulation that changes the phase of the carrier wave [5][6]. The principle of modulation is a phase rotation of 180 degrees occurring for the logical unit. For logical zero, the phase does not rotate, it remains the same as for the previous one symbol. The main advantage of DPSK is the intensity improvement of 3 dB, which means better sensitivity of the receiver compared to OOK modulation. Fig.1 shows DPSK modulation principle - for each logical unit there is a phase shift of 180°. Logic zero does not shift the phase [7][8].

Differential Quadrature Phase Shift Keying (DQPSK) has gained popularity due to its ability to transmit signals over long distances. In addition, it is less prone to noise when passing through several narrow channels in wavelength division multiplexing (DWDM) [5]. In the case of DQPSK modulation, the bit pairs correspond to the given phase shifts (not specific phase values). A phase shift from the initial phase is given or we can consider a 90° phase shift between adjacent symbols.

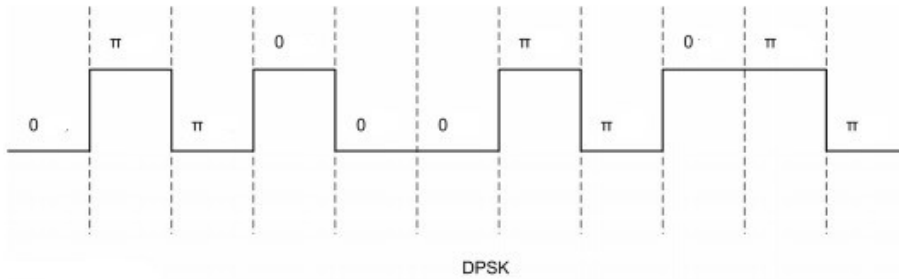


Fig. 1 DPSK modulation principle

### III. DESIGN OF ULTRA-DWDM SYSTEM WITH DIFFERENT OPTICAL MODULATIONS

The design consists of lasers, modulators, a multiplexer and a demultiplexer, an optical fiber, a dispersion compensating optical fiber (DCF), Erbium doped fiber amplifiers (EDFA) and optical receivers. The parameters of the 30-channel DWDM of the proposed model are the transmission speed of 40 Gb/s, the central frequency is 193,1 THz, the frequency spacings are 100 GHz. The total optical distance is 60 km. The DWDM design is the same for both modulations.

In picture Fig. 2 the first block on the left contains 30 lasers. The second block contains the modulations shown in the Fig. 4. In Fig. 3 is receiving part of DWDM scheme from OptiSystem.

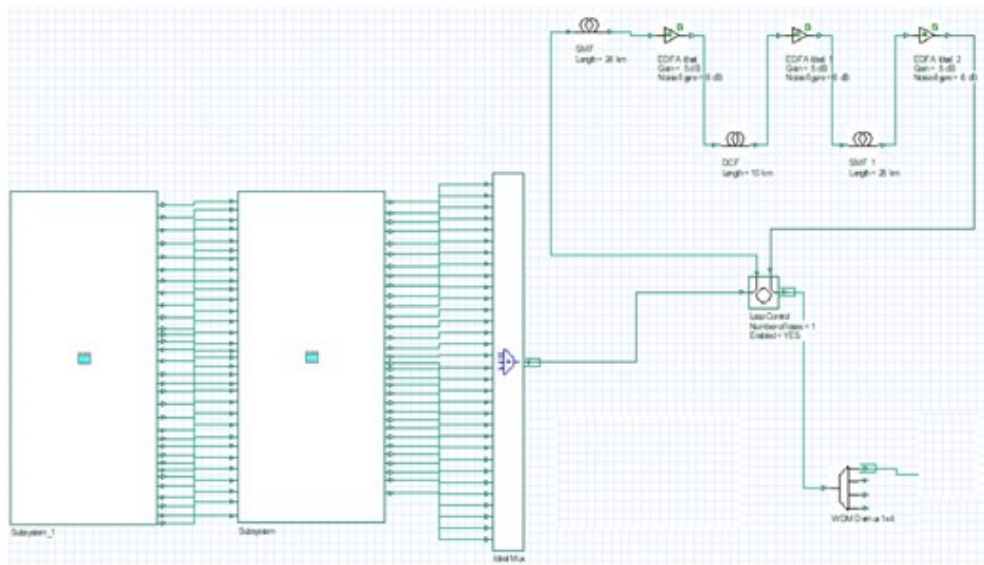


Fig. 2 Schematic of a DWDM communication system at a transmission speed of 40 Gb/s

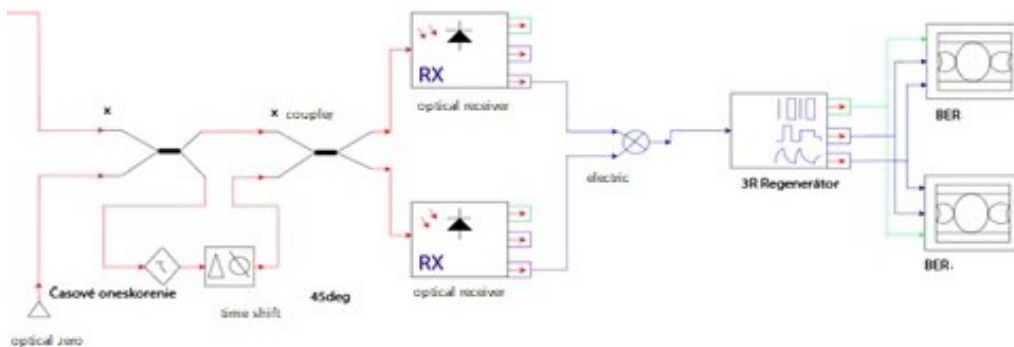


Fig. 3 Receiving side of DWDM scheme from OptiSystem



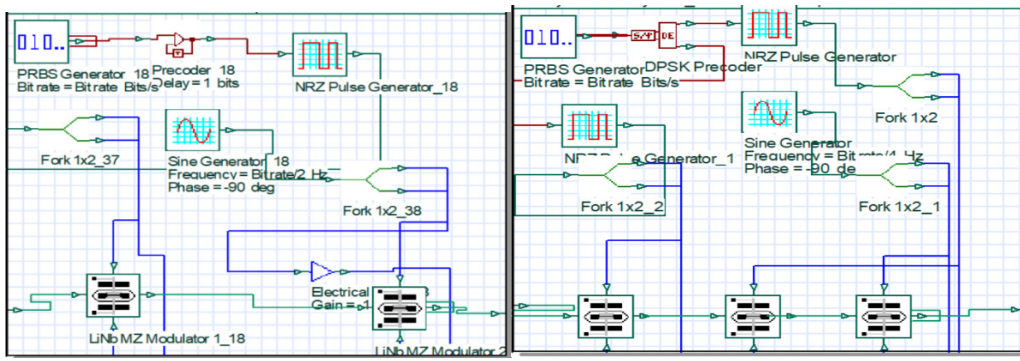


Fig. 4 On the left side OptiSystem scheme DPSK, on the right side OptiSystem scheme DQPSK

From the eye diagram we can extract various parameters of received signal. The most used ones are bit error rate (BER), Q factor measurements measured in decibels and the eye opening. Eye opening signals how much is the analyzed eye diagram opened. The most open part means the best signal-to-noise ratio (SNR). Q factor is a function of SNR and provides a qualitative description of receiver performance. OptiSystem allows to plot BER and Q factor separately (Fig.5, Fig.6).

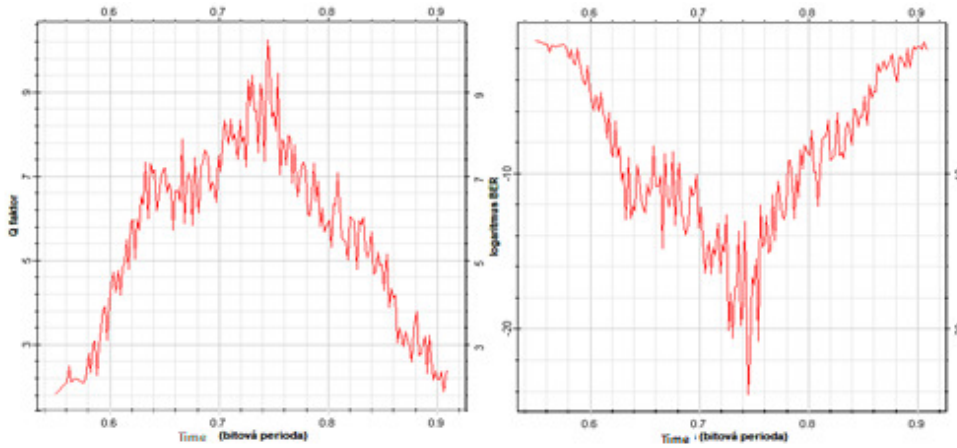


Fig. 5 Q factor on the left side, BER on the right side for DPSK modulation

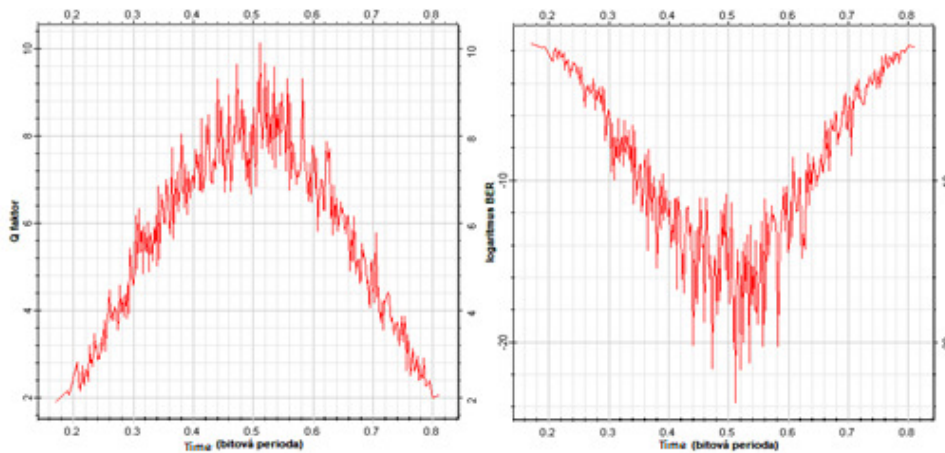


Fig. 6 Q factor on the left side, BER on the right side for DQPSK modulation

Our presumption is that the performance of the system with DPSK modulation will be better, according to the comparison of the results from our simulations. The simulation is performed with components with ideal parameters and without negative effects on the system, such as nonlinear phenomena. The aim of the work was to compare DPSK and DQPSK modulation and not the effects acting on the optical network. In real conditions, the results will not be as good as in the simulation (Tab.1).

Tab. 1 Parameters of modulations in DWDM

<b>DPSK modulation</b>	
Frequency	193,1 THz
Wavelength	1552,5244 nm
BER	$6,19 \times 10^{-25}$
Q factor	10,2448
Eye height	0,0009
<b>DQPSK modulation</b>	
Frequency	194,0 THz
Wavelength	1545,3219 nm
BER	$1,8369 \times 10^{-24}$
Q factor	10,1398
Eye height	0,0007

#### IV. CONCLUSION

The use of optical communication is an essential part of human society. Changing important parameters such as path length, baud rate, laser gain, frequency change can have a positive or negative effect on the received signal for modulation. When examining modulations, it can be determined that every single modulation has its positives and negatives. The DWDM system is very well optimized for 40 Gb/s. By changing the parameters on the transmitting, receiving or channel part, OptiSystem may not draw the correct BER eye. Multistate modulations (with more than two states), for example DPSK, DQPSK, combine phase and amplitude modulation or phase and polarization. These modulations appear to be promising for high-speed and long-distance transmission systems. In the future, we can count on higher transmission speeds and we can also rely on these modulation formats to be or are already adapted to the everyday requirements of the human being.

#### ACKNOWLEDGMENT

This work was supported by research projects VEGA 1/0584/20 “UWB sensor systems for monitoring people working in real conditions (ReConLoc)” and FEEI-2021-71 “Methods of data security in modern communication systems”.

#### REFERENCES

- [1] HUI, Rongqing. Introduction to Fiber-Optic Communications. Cambridge, Massachusetts, USA: Academic Press, 2020. ISBN 978-0-12-805345-4.
- [2] Biswanath Mukherjee, Ioannis Tomkos, Peter Winzer, Yongli Zhao: Handbook of Optical Networks. Switzerland: Gewerbestrasse 11, 6330 Cham, 2020, ISBN: 978-3-030-16250-4, 1169 pp.
- [3] Agrawal, Govind P.: Fiber-Optic Communication Systems (Third Edition). New York: The Institute of Optics University of Rochester, 2002, ISBN: 0-471-22114-7, 546 pp.
- [4] Winzer, Peter J. Energy-Efficient Optical Transport Capacity Scaling Through Spatial Multiplexing. IEEE Photonics Technology Letters. 2011, 23(13), 851 - 853. ISSN 1041-1135. Available on: doi:10.1109/LPT.2011.2140103
- [5] Huszanik, Tomáš, Ján Turán and Luboš Ovseník. On the Impact of Fiber Nonlinear Effects on the CP-DQPSK Modulated Ultra-DWDM System. Acta Electrotechnica et Informatica. 2019, 19(4), 21-28. ISSN 1335-8243. Available on: doi:10.15546/aei-2019-0026
- [6] G. Bosco, A. Carena, Member, IEEE, V. Curri, R. Gaudino and P. Poggiolini, —On the Use of NRZ, RZ, and CSRZ Modulation at 40 Gb/s With Narrow DWDM Channel Spacing, IEEE Journal of Lightwave Technology, Vol. 20, No. 9, Pages 1694-1704, 2002.
- [7] G.694.1: Spectral grids for WDM applications: DWDM frequency grid. ITU-T, 2020. Available on: <https://www.itu.int/rec/T-REC-G.694.1/en>
- [8] Laude, J.Pierre. Dwdm Fundamentals, Components, and Applications: Artech House Optoelectronics Library. London: Artech House, 2002. ISBN 978-1580531771.

# Dynamická identifikácia vnútorných veličín asynchrónneho motora

<sup>1</sup>Marek FEDOR, <sup>2</sup>Daniela PERDUKOVÁ

<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>marek.fedor@student.tuke.sk, <sup>2</sup>daniela.perdukova@tuke.sk

**Abstract** — V článku je navrhnutá nová identifikačná metóda problémovo merateľných vnútorných veličín asynchrónneho motora (AM), ako sú zložky vektora magnetického toku a elektromagnetický moment. Pre túto identifikáciu sú použité bežne merateľné veličiny AM, ako sú prúdy statora, frekvencia statorového napätia a mechanická uhlová rýchlosť, na základe ktorých sa zisťuje spätný vplyv vektora rotorového toku na vektor statorových prúdov AM. Na základe tohto spätného vplyvu je možné identifikovať aj aktuálnu hodnotu rotorového odporu, ktorá sa počas prevádzky AM mení a má výrazný vplyv na výslednú kvalitu identifikovaných veličín, ako aj nadradeného riadenia AM. Stabilita identifikačnej štruktúry je zaručená koreňmi charakteristickej rovnice jej lineárneho prenosu. Simulačné výsledky potvrdzujú kvalitu, efektívnosť a robustnosť navrhnujej identifikačnej metódy.

**Keywords** — asynchrónny motor, magnetický tok, moment motora, identifikácia systému

## I. ÚVOD

Použitie asynchrónnych motorov v elektrických pohonoch má svoje nesporné výhody, ako sú jednoduchá konštrukcia, údržba, spoľahlivosť, účinnosť atď. V priemyselnej praxi sa asynchrónne motory používajú pre konvenčné, ale aj výkonovo náročné aplikácie [1]-[4]. Z hľadiska riadenia sú to voči jednosmerným cudzobudeným motorom podstatne zložitejšie regulované systavy, ktorých riadenie vyžaduje komplikovanejšie akčné členy a vo všeobecnosti podstatne zložitejšie algoritmy riadenia [5]-[7]. V súčasnosti sú už dobre prepracované frekvenčné meniče, ktoré je možné použiť ako akčné členy pre pohony s asynchrónnymi motormi. Pre kvalitné algoritmy riadenia je potrebné identifikovať dôležité parametre a vnútorné veličiny stroja, ktoré sa počas prevádzky systémov s elektrickými pohonmi môžu meniť napríklad vplyvom zmeny teploty alebo nasýtenia magnetického toku.

V literatúre sa uvádzajú viaceré prístupy k určovaniu základných stacionárnych parametrov asynchrónnych motorov, ako sú odpory vinutí, vlastné a vzájomné indukčnosti (príp. hlavná a rozptylová indukčnosť) a moment zotrvačnosti na hriadeli motora, ktorých presné určenie môže výrazne ovplyvniť jeho účinnosť. V literatúre [8] je použitá metóda konečných prvkov pre identifikáciu parametrov asynchrónneho motora a jeho dynamického správania. Prehľad metód identifikácie statorového a rotorového odporu pri bezsnímačovom riadení AM a návrh jeho pozorovateľa je uvedený v literatúre [9]. Prehľadový článok o metódach on-line a off-line identifikácie parametrov AM je uvedený v literatúre [10].

Pre kvalitné dynamické riadenie asynchrónnych motorov je nevyhnutné okrem poznania základných stacionárnych parametrov AM identifikovať v reálnom čase dôležité vnútorné veličiny stroja, ako sú rotorové toky a z nich odvodený vnútorný dynamický moment motora na hriadeli rotora. Pretože priame meranie týchto veličín je náročné a často aj z hľadiska potrieb praxe nemožné, získavajú sa okamžité hodnoty týchto veličín nepriamo pomocou rôznych, viac alebo menej náročných a presných identifikačných metód. Presnosť identifikácie u týchto metód je väčšinou závislá od presnej znalosti rotorového odporu motora, resp. jeho rotorovej časovej konštanty. V literatúre [11] je prezentovaný spôsob identifikácie rotorového toku na báze neurónových sietí a referenčného modelu stroja, ktorý pracuje aj pri nízkych otáčkach AM. On-line identifikácia statorového odporu, rotorového odporu a záťažného momentu stroja je prezentovaná v literatúre [12]. Kvalitná identifikácia vnútorných veličín AM je závislá od presnej identifikácie parametrov stroja, ktorá je zase pri metódach s referenčným modelom závislá hlavne pri bezsnímačovom riadení na presných hodnotách identifikácie dynamických veličín stroja. Tento problém je riešený v literatúre [13] návrhom pozorovateľa, ktorý sa snaží

súčasne optimalizovať obidva tieto identifikačné procesy.

V článku je navrhnutá nová dynamická identifikačná štruktúra pre AM vychádzajúca z jeho matematického modelu, ktorá umožňuje súčasne kvalitne identifikovať rotorový tok motora a jeho moment, ako aj premenlivý rotorový odpor. Táto identifikačná metóda je založená na určení spätného vplyvu vektora rotorového toku na vektor statorových prúdov AM, pričom je potrebná znalosť len dostupne merateľných veličín AM ako sú jeho statorové prúdy a mechanická uhlová rýchlosť. Stabilita identifikácie tohto spätného vplyvu je zaručená koreňmi charakteristickej rovnice prenosu lineárneho identifikačného obvodu. Celá identifikácia prebieha on-line počas štandardnej prevádzky pohonu. Dosažené výsledky potvrdzujúce efektívnosť a kvalitu navrhutej identifikačnej metódy boli overené simulačnými meraniami.

## II. MATEMATICKÝ MODEL ASYNCHRÓNNEHO MOTORA

V literatúre existuje mnoho rôznych typov matematických modelov asynchrónneho motora podľa toho, ktoré veličiny motora sú zvolené za jeho stavové veličiny. Všeobecne môžeme povedať, že asynchrónny motor je sústava piatich nelineárnych diferenciálnych rovníc prvého rádu, ktorú nevieme analyticky riešiť. Pri voľbe prúdu statora a toku rotora za stavové veličiny modelu môžeme popísať asynchrónny motor nakrátko nasledujúcimi rovnicami [14]

$$\begin{bmatrix} \frac{di_{1x}}{dt} \\ \frac{di_{1y}}{dt} \\ \frac{d\psi_{2x}}{dt} \\ \frac{d\psi_{2y}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega_0 & \omega_1 & -K_{12}\omega_g & -K_{12}\omega_m n_p \\ -\omega_1 & \omega_0 & K_{12}\omega_m n_p & -K_{12}\omega_g \\ M\omega_g & 0 & -\omega_g & \omega_2 \\ 0 & M\omega_g & -\omega_2 & \omega_g \end{bmatrix} \begin{bmatrix} i_{1x} \\ i_{1y} \\ \psi_{2x} \\ \psi_{2y} \end{bmatrix} + \begin{bmatrix} K_{11} & 0 \\ 0 & K_{11} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1x} \\ u_{1y} \end{bmatrix} \quad (1)$$

$$n_p \frac{M}{L_2} (\psi_{2x} i_{1y} - \psi_{2y} i_{1x}) - T_{load} = J \frac{d\omega_m}{dt} \quad (2)$$

V tomto modeli sú jednotlivé veličiny motora (vektor statorového prúdu  $\mathbf{i}_1$  a vektor rotorového toku  $\boldsymbol{\psi}_2$ ) vyjadrené pomocou ich zložiek v pravouhlej súradnicovej sústave  $x$ - $y$ , ktorá rotuje voči statoru uhlovou rýchlosťou statorového poľa  $\omega_1$ . Jednotlivé parametre v rovniciach (1-2) vieme určiť z parametrov motora prepočítaním podľa nasledujúcich vzťahov

$$K_{11} = \frac{3}{2} \left( L_{s1} + \frac{L_{s2} L_m}{L_{s2} + L_m} \right)^{-1} \quad (3)$$

$$K_{12} = -\frac{3}{2} \left( L_{s1} + L_{s2} + \frac{L_{s1} L_{s2}}{L_m} \right)^{-1} \quad (4)$$

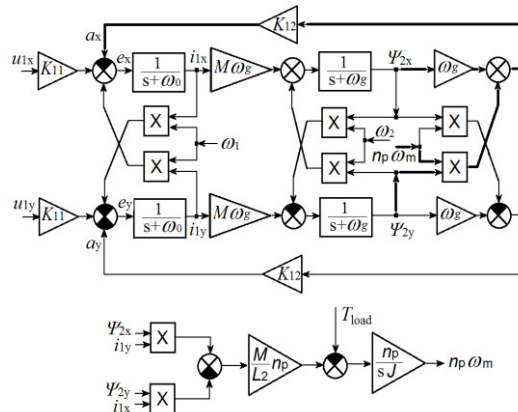
$$\omega_0 = K_{11} \left[ R_1 + \left( \frac{M}{L_2} \right)^2 R_2 \right] \quad (5)$$

$$M = \frac{2}{3} L_m \quad (6)$$

$$\omega_g = \frac{R_2}{L_2} \quad (7)$$

$$L_2 = \frac{2}{3} (L_{s2} + L_m) \quad (8)$$

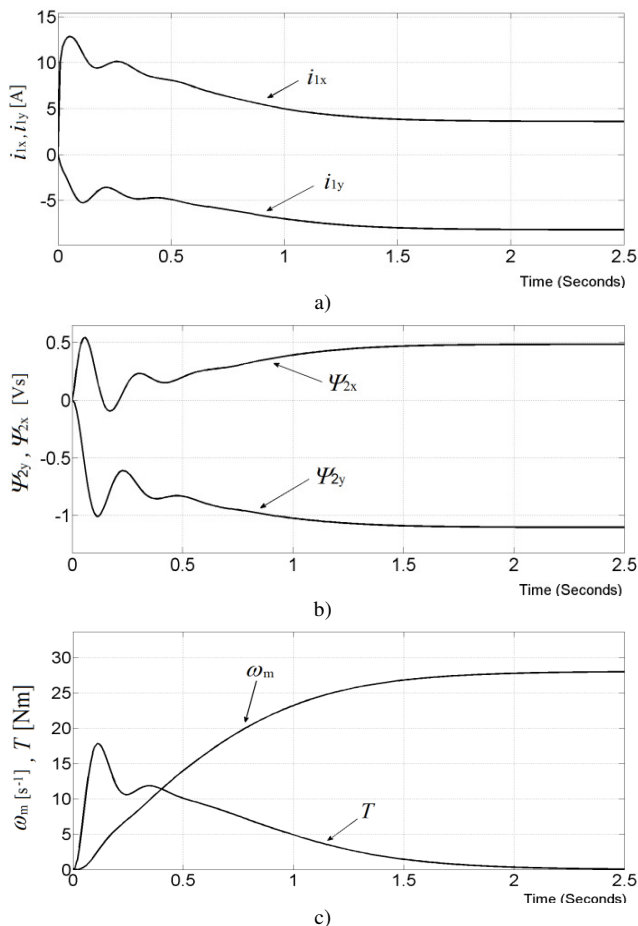
Bloková schéma asynchrónneho motora zodpovedajúca rovniciam (1-2) s parametrami upravenými podľa uvedených vzťahov (3-8) je na Obr. 1.



Obr.1 Bloková schéma asynchrónneho motora v súradnicovej sústave  $x$ - $y$  zviazanej so statorovým poľom

Dynamika jednotlivých veličín motora (statorové prúdy, rotorové toky a uhlová rýchlosť) po pripojení na statorové napätie  $U_1=40V$  s frekvenciou  $\omega_1=28,03$  rad/s získaná číslcovou simuláciou je uvedená na Obr.2.

Pre overovanie dosiahnutých výsledkov číslcovou simuláciou bol v článku použitý program MATLAB/Simulink. Parametre AM motora pre všetky simulácie sú uvedené v Prílohe. Pri overovaní vlastností číslcovou simuláciou ďalej predpokladáme, že asynchrónny motor je prevádzkovaný pri skalárnom riadení napätím o veľkosti  $U_1$  pričom platí pre jeho frekvenciu známy vzťah  $U_1/\omega_1=const$ . V tomto prípade môžeme považovať  $x$ -ovú zložku napätia  $u_{1x}$  za totožnú s veľkosťou statorového napätia  $U_1$ , t.j.  $U_1=u_{1x}$ .



Obr.2 Dynamické odozvy a) statorových prúdov, b) rotorových tokov a c) uhlovej rýchlosti AM na skok vstupného napätia a frekvencie

### III. NÁVRH IDENTIFIKAČNEJ ŠTRUKTÚRY SPÄTNÉHO VPLYVU ROTORA NA STATOROVÉ PRÚDY

Základom pre identifikáciu ťažko merateľných „vnútorných“ veličín AM ako je vektor rotorového toku a elektromagnetický moment je čo najkvalitnejšia identifikácia spätného vplyvu vektora rotorového toku na vektor statorového prúdu. V blokovej schéme na Obr.1 je tento vplyv označený veličinou  $a_x$ , resp.  $a_y$ . Cieľom identifikácie „vnútorných“ veličín AM je čo najpresnejšie určenie ich okamžitých hodnôt.

Predpokladajme, že na danom AM vieme merať statorové prúdy  $i_{1x}$ ,  $i_{1y}$ , mechanickú uhlovú rýchlosť  $\omega_m$  a poznáme okamžité hodnoty vstupných veličín statora  $U_1$  a  $\omega_1$ . Tieto veličiny sú obvykle dobre merateľné. Pre zložku prúdu  $i_{1x}$  statorového prúdu z prvého riadku rovnice (1) platí, že okamžitá hodnota spätného vplyvu je

$$a_x = K_{12}(\omega_g \psi_{2x} + \omega_m n_p \psi_{2y}) \quad (9)$$

a podobne pre zložku statorového prúdu  $i_{1y}$  platí

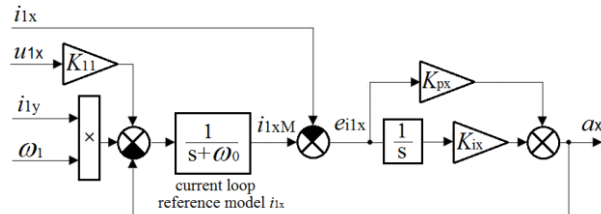
$$a_y = K_{12}(\omega_g \psi_{2y} - \omega_m n_p \psi_{2x}) \quad (10)$$

Pre kvalitnú dynamickú identifikáciu týchto veličín bola navrhnutá identifikačná štruktúra podľa Obr.3, kde je zobrazená len identifikácia zložky  $a_x$  (identifikovaná hodnota je označená ako  $a_x^*$ ). Je to štruktúra s referenčným modelom statorového prúdového obvodu motora, v ktorej sa do referenčného modelu privádza ten istý vstupný signál, aký pôsobí na jednotlivú

zložku statorového prúdu. Spätňý vplyv rotora na túto zložku prúdu statora je identifikovaný z odchýlky nameraného prúdu  $i_{1x}$  a jeho predpokladanej hodnoty z referenčného modelu  $i_{1xM}$ . Táto odchýlka je spracovávaná PI obvodom, ktorého výstup  $a_x^*$  sa nastaví (identifikuje) na hodnotu  $a_x$  z motora. Laplaceov prenos tohto lineárneho obvodu je

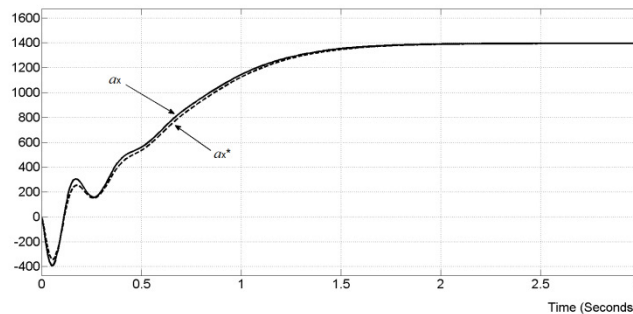
$$F(s) = \frac{\frac{K_{px}s + K_{ix}}{s}}{1 + \frac{K_{px}s + K_{ix}}{s} \frac{1}{s + \omega_0}} = \frac{K_{px}s^2 + (K_{px}\omega_0 + K_{ix})s + K_{ix}\omega_0}{s^2 + (K_{px} + \omega_0)s + K_{ix}} \quad (11)$$

a jeho stabilitu, ako aj dynamiku vieme jednoducho nastaviť vhodnou voľbou zosilnení jeho proporcionálnej zložky  $K_{px}=10$  a integračnej zložky  $K_{ix}=11870$ . Identifikačná štruktúra na Obr.3 je lineárna a preto je jej stabilita zaručená zápornou reálnou zložkou pólov menovateľa a prenosu z rovnice (11).

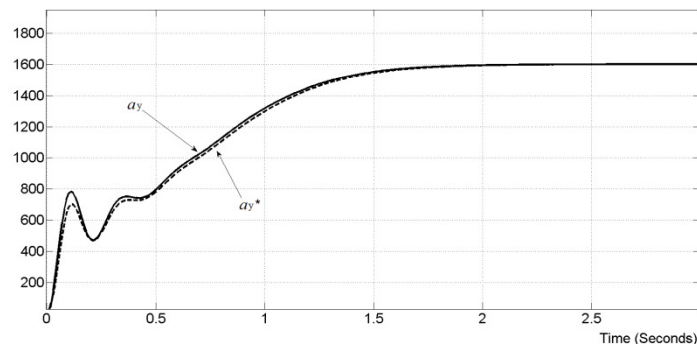


Obr.3 Štruktúra identifikácie spätňého vplyvu rotora na stator  $a_x^*$

Analogicky bola navrhnutá aj identifikačná štruktúra pre identifikáciu spätňého vplyvu rotora pre zložku  $a_y$ . Vlastnosti navrhutej identifikačnej štruktúry boli overené číslicovou simuláciou. Na Obr.4 sú zobrazené priebehy reálnych “vnútorných” veličín  $a_x$  a  $a_y$  z AM, ako aj ich identifikované hodnoty  $a_x^*$  a  $a_y^*$ . Z uvedených priebehov je zrejmé, že navrhnutá identifikačná štruktúra vykazuje výborné dynamické vlastnosti, pričom je principiálne stabilná. Vzhľadom na presnosť snímania prúdov a napätí, ktoré sa pohybujú rádovo v percentách, môžeme odchýlku medzi reálnymi a identifikovanými hodnotami považovať za zanedbateľnú predovšetkým v dynamických stavoch, pretože v ustálenom stave je táto odchýlka nulová.



a)



b)

Obr.4 Kvalita identifikácie veličín a)  $a_x^*$ , b)  $a_y^*$

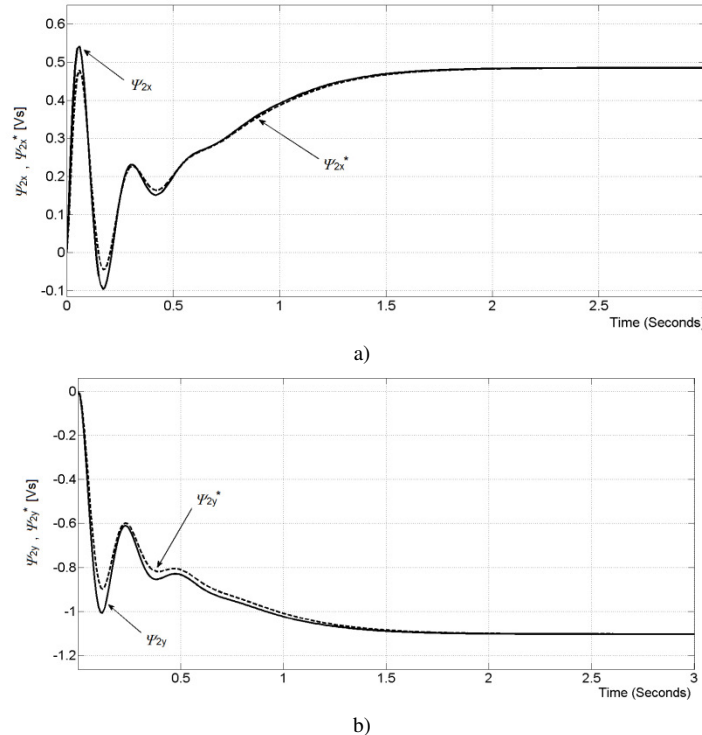
#### IV. IDENTIFIKÁCIA TOKOV ROTORA A MOMENTU ASYNCHRÓNNEHO MOTORA

Ak vieme kvalitne identifikovať spätňý vplyv vektora rotorových tokov na vektor statorových prúdov AM pomocou veličín  $a_x^*$  a  $a_y^*$  na základe navrhutej identifikačnej štruktúry z Obr.3, potom môžeme určiť zložky rotorového toku AM z blokovej schémy na Obr.1 podľa nasledujúcich rovníc

$$\Psi_{2x}^* = \frac{1}{K_{12}} \left( \frac{\omega_g a_x^* - \omega_m a_y^*}{\omega_g^2 + \omega_m^2} \right) \quad (12)$$

$$\Psi_{2y}^* = \frac{1}{K_{12}} \left( \frac{\omega_g a_y^* + \omega_m a_x^*}{\omega_g^2 + \omega_m^2} \right) \quad (13)$$

Kvalita identifikácie jednotlivých zložiek rotorového toku v porovnaní s ich reálnymi hodnotami získaná simulačnými meraniami je ukázaná na Obr.5.



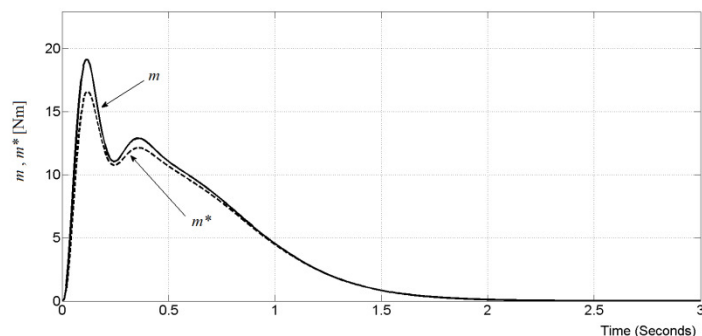
Obr.5 Porovnanie identifikovaných a skutočných zložiek rotorových tokov AM a)  $\Psi_{2x}$ ; b)  $\Psi_{2y}$

Z obrázka 5 vidíme, že hodnoty jednotlivých zložiek vektora rotorového toku AM sa takmer zhodujú s jeho identifikovanými hodnotami. Najväčšia odchýlka v dynamickom stave je približne v čase  $t=0.15$  s, pričom jej hodnota je menšia ako 4%, čo je pre použitie v praktických radiaciach aplikáciách úplne vyhovujúce. V ustálenom stave je identifikačná odchýlka jednotlivých zložiek rotorového toku nulová, čo odpovedá nulovej identifikačnej odchýlke veličín  $a_x^*$  a  $a_y^*$  na Obr.4.

Opäť platí, že ak vieme kvalitne identifikovať jednotlivé zložky rotorového toku AM, potom jeho elektromagnetický moment môžeme určiť na základe rovnice (2) nasledovne

$$m^* = \frac{Mn_p}{L_2} (\Psi_{2x}^* i_{1y} - \Psi_{2y}^* i_{1x}) \quad (14)$$

Výsledný priebeh identifikácie elektromagnetického momentu získaný číslicovou simuláciou na základe rovnice (14) je uvedený na Obr.6. Tento priebeh potvrdzuje kvalitu identifikácie v ustálenom aj dynamickom stave. V ustálenom stave je identifikačná odchýlka opäť nulová. Najväčšia chyba identifikácie v dynamickom stave je v čase  $t=0.15$  s a predstavuje približne 3 % z maximálneho rozsahu priebehu elektromagnetického momentu AM.

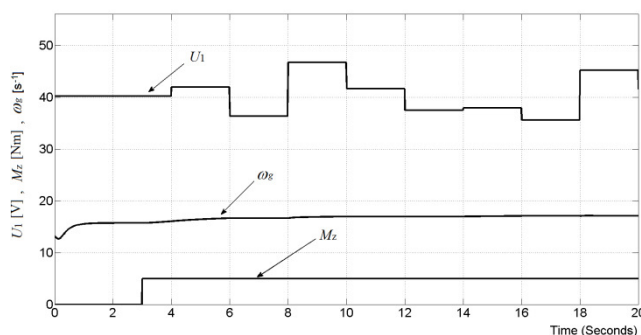


Obr.6 Porovnanie identifikovaného elektromagnetického momentu  $m^*$  a skutočného momentu  $m$  AM

V. VERIFIKÁCIA IDENTIFIKÁCIE ROTOROVÉHO TOKU A MOMENTU ASYNCHRÓNNEHO MOTORA  
 PRE RÔZNE PREVÁDZKOVÉ STAVY MOTORA

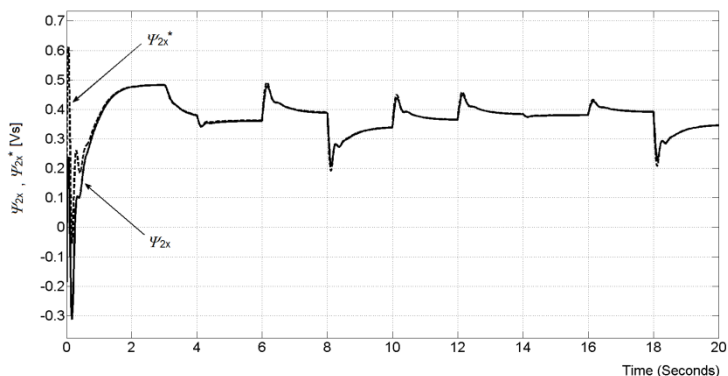
Vlastnosti metódy identifikácie rotorového toku a momentu AM popísané v Kap.2 a Kap.3 boli overované pre rôzne prevádzkové stavy AM číslcovou simuláciou.

Predpokladajme, že asynchrónny motor je prevádzkovaný pri skalárnom riadení napätím o veľkosti  $U_1$  (t.j.  $U_1=u_{1x}$ ) a pri zmenách záťažného momentu  $M_z$ . Z uvedeného dôvodu boli simulované kvázi náhodné skokové zmeny vstupného napätia  $U_1$  na statore AM motora a skoková zmena jeho záťažného momentu  $M_z$ , tak, ako je to ukázané na Obr.10. V čase  $t=3$  s bol na motor privedený skok záťažného momentu o veľkosti  $M_z=5$  Nm. Skoky veľkosti a frekvencie satorového napätia  $U_1$  sa generujú náhodne v dvojsekundových časových intervaloch, pričom je dodržaný vzťah  $U_1/\omega_1=\text{konšt.}$ , ako je to bežné u štandardných meničov. Uvedenú prevádzku je možné považovať za najhoršie prípady pre identifikáciu rotorového odporu, resp. jemu priamo úmerného parametra  $\omega_g$ , ktorý sa v tomto prípade adaptoval z počiatočnej hodnoty  $\omega_{g0}=13,18$  s<sup>-1</sup> na hodnotu  $\omega_g=17,18$  s<sup>-1</sup> tak, ako je to ukázané na Obr.7.

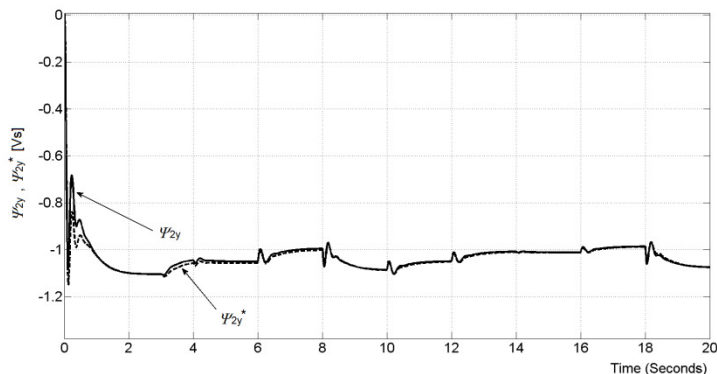


Obr. 7 Priebehy vstupných veličín  $U_1$ ,  $M_z$  a priebeh adaptácie parametra  $\omega_g$

Obrázok 11 ukazuje, ako sa zlepšuje kvalita identifikácie jednotlivých zložiek rotorového toku AM a identifikácia jeho elektromagnetického momentu pri zohľadnení adaptácie hodnoty parametra  $\omega_g$  z počiatočnej hodnoty  $\omega_{g0}=13,18$  s<sup>-1</sup> na hodnotu  $\omega_g=17,18$  s<sup>-1</sup>. Pretože uvažujeme rozbeh AM z nenabudeného stavu, v prvej sekunde rozbehu počas nabudzovania je chyba identifikácie najvýraznejšia. Po nabudení AM sa chyba dostane prakticky do rozsahu  $\pm 1\%$  z aktuálnej hodnoty v ustálených a aj počas dynamických stavov jednotlivých identifikovaných veličín tak, ako je to evidentné z Obr.8.

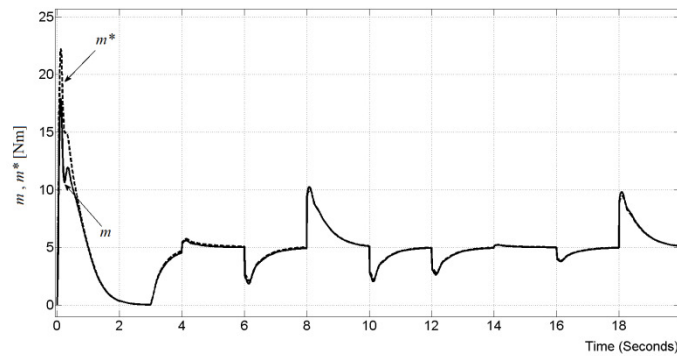


a)



b)





c)

Obr.8 Priebeh identifikácie jednotlivých zložiek rotorového toku a)  $\Psi_{2x}$ ; b)  $\Psi_{2y}$  a c) momentu motora  $m$  pri zohľadnení adaptácie rotorového odporu, resp. parametra  $\omega_g$

## VI. ZÁVER

V predloženom článku je ukázaný návrh originálnej metódy kvalitnej identifikácie rotorových tokov a elektromagnetického momentu AM založený na identifikácii spätného vplyvu rotora na stator, ktorá používa pre svoju činnosť bežne dostupné merateľné veličiny AM, a to prúdy statora a mechanickú uhlovú rýchlosť rotora. Identifikačná štruktúra pre identifikáciu spätného vplyvu rotora na stator AM je lineárna a jej stabilita je zaručená koreňmi charakteristickej rovnice jej prenosu. Navrhnutý postup identifikácie je možné rozdeliť do nasledujúcich krokov:

1. Identifikácia spätného vplyvu rotora (presnejšie rotorového toku) na stator (presnejšie statorový prúd) AM, t.j. identifikácia veličín  $a_x$  a  $a_y$ .
2. Výpočet zložiek rotorového toku z mechanickej uhlovej rýchlosti rotora a z identifikovaných veličín spätného vplyvu rotora na stator –  $a_x^*$  a  $a_y^*$ .
3. Výpočet elektromagnetického momentu motora z meraných statorových prúdov a identifikovaného vektora rotorového toku motora.

Kvalita a efektívnosť navrhnutej identifikačnej metódy bola overená číslcovou simuláciou v programe MATLAB/Simulink.

Výsledky získané číslcovou simuláciou (Fig.4-Fig.6) ukazujú, že navrhnutá metóda dokáže s dostatočnou presnosťou v ustálených a dynamických stavoch identifikovať spätný vplyv rotora na stator AM motora, a rovnako potom jeho vnútorné rotorové magnetické toky a elektromagnetický moment.

Identifikované vnútorné veličiny motora, ako sú jednotlivé zložky jeho rotorového toku a vnútorný elektromagnetický moment motora, je možné použiť pre rôzne v praxi bežne používané metódy dynamickej regulácie pohonu s takýmto strojom, ako sú skalárne riadenie, vektorové riadenie rôzneho typu a priame riadenie momentu AM.

## APPENDIX

$P_N=3$  KW,  $U_{1N}=220$  V,  $I_{1N}=6.9$ ,  $m_N=20$ Nm,  $n_N=1430$  otmin<sup>-1</sup>,  $n_p=2$ ,  $J=0.1$  kgm<sup>2</sup>,  $R_1=1.8$   $\Omega$ ,  $R_2=1.85$   $\Omega$ ,  $L_h=0.202$  H,  $L_{s1}=L_{s2}=0,0086$  H,  $K_{11}=59.3514$  H<sup>-1</sup>,  $K_{12}=-56.9277$  H<sup>-1</sup>,  $\omega_g=13.177$  s<sup>-1</sup>,  $M=0.1347$  H,  $\omega_0=207.9$  s<sup>-1</sup>,  $L_2=0,1404$  HACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-19-0210.

## REFERENCES

- [1] Y. Liu, J. Zhao, R. Wang, and C. Huang, "Performance improvement of induction motor current controllers in field-weakening region for electric vehicles," *IEEE Trans. Power Electron.*, vol. 28, no. 5, 2013, pp. 2468–2482.
- [2] C. Sun, S. Ai, L. Hu, and Y. Chen, "The development of a 20 MW PWM driver for advanced fifteen-phase propulsion induction motors," *J. Power Electron.*, vol. 15, no. 1, 2015, pp. 146–159.
- [3] M. H. Holakoote, A. Taheri, and M. B. B. Sharifian, "MRAS based speed estimator for sensorless vector control of a linear induction motor with improved adaptation mechanisms," *J. Power Electron.*, vol. 15, no. 5, 2015, pp. 1274–1285.
- [4] J. Vittek, P. Butko, B. Ftorek, P., L. Gorel, „Energy Near-Optimal Control Strategies for Industrial and Traction Drives with a.c. Motors.“. *Mathematical Problems in Engineering*, Volume 2017, Hindawi Publishing Corporation, 22 pages. ID 1857186. <https://doi.org/10.1155/2017/1857186>
- [5] B.K Bose, *Modern Power Electronics and AC Drives*; Prentice-Hall: Englewood Cliffs, NJ, USA, 2002; ISBN 0-13-016743-6.
- [6] M.P Kazmierkowski, A.B. Kasprowicz, „Improved direct torque and flux vector control of PWM inverter-fed induction motor drives“, *IEEE Trans. Ind. Electron.*, 42, 1995, pp. 344–350.
- [7] Y. Zhang, Y. Bai, H. Yang, „A Universal Multiple-Vector-Based Model Predictive Control of Induction Motor Drives“, *IEEE Trans. Power Electron.*, 33, 2018, pp. 6957–6969.

- [8] S. Williamson, L. H. Lim and A. C. Smith, „Transient analysis of cage induction motor using finite elements,“ *IEEE Transactions on Magnetics*, Vol. 26, No.2, 1990, pp. 941-944.
- [9] J. Chen, J. Huang, „Stable Simultaneous Stator and Rotor Resistances Identification for Speed Sensorless IM Drives: Review and New Results.“ *IEEE Transactions on Power Electronics*, Vol. 33, No. 10, 2018.
- [10] C. M. Verrelli, A. Savoja, M. Mengoni, R. Marino, P. Tomei, L. Zarri, „On-Line Identification of Winding Resistances and Load Torque in Induction Machines“, *IEEE Transactions on Control Systems Technology*, Vol. 22, No. 4, 2014.
- [11] G. Sheng-wei, „Stator flux linkage observer based on RBF neural network for IM,“ in *Third Int. Conf. on Intelligent Networks and Intelligent Systems*, Shenyang, China, 2010. DOI: 10.1109/ICINIS.2010.81
- [12] K.-Y. Chen, W.-H. Yang, and R.-F. Fung, „System identification by using RGA with a reduced-order robust observer for an induction motor,“ *Mechatronics*, vol. 54, pp. 1-15, 2018. DOI: 10.1016/j.mechatronics.2018.07.00
- [13] K.-Y. Chen, W.-H. Yang, and R.-F. Fung, „System identification by using RGA with a reduced-order robust observer for an induction motor,“ *Mechatronics*, vol. 54, pp. 1-15, 2018. DOI: 10.1016/j.mechatronics.2018.07.001
- [14] K.-Y. Chen, W.-H. Yang, and R.-F. Fung, „System identification by using RGA with a reduced-order robust observer for an induction motor,“ *Mechatronics*, vol. 54, pp. 1-15, 2018. DOI: 10.1016/j.mechatronics.2018.07.001

# Využitie technológie Blockchain v 6G a MANET sieti

<sup>1</sup>Richard NEZNIK, <sup>2</sup>Jan PAPAJ

<sup>1,2</sup>Katedra Elektroniky a Multimediálnych Telekomunikácií,  
Fakulta Elektrotechniky a Informatiky Technická Univerzita Košice,  
Slovenská Republika

<sup>1</sup>richard.neznik@student.tuke.sk, <sup>2</sup>jan.papaj@tuke.sk

**Abstrakt** – Článok poskytuje úvodnú štúdiu využitia technológie Blockchain pre viacpreskokové mobilné siete, ako aj jej implementáciu do mobilných sieťach 6. generácie. Cieľom článku je poskytnúť základný popis využitia technológie blockchain s dôrazom kladeným na zvýšenie sieťovej bezpečnosti v mobilných sieťach, ako aj využitie jeho potenciálu v procese smerovania IoT dát vo viacpreskokových sieťach typu MANET, DTN, DRONET a FLYNET.

**Kľúčové slová** – 6G siete, Bezpečnosť, Blockchain, MANET, DTN, DRONET.

## I. ÚVOD

Nedávny rozmach diverzifikovaných mobilných aplikácií, najmä tých, ktoré podporujú umelú inteligenciu (Artificial Intelligence – AI), podnietili búrlivé diskusie o budúcom vývoji bezdrôtovej komunikácie. Piata generácia mobilných sietí sa už v niektorých častiach sveta zavádza s očakávaním riešenia obmedzení pre súčasné celulárne systémy a poskytovania základnej platformy pre spravodajské služby, ktoré sa budú rozvíjať a prosperovať.

Dnes žijeme v dobe, ktorú môžeme nazvať ako plne digitálnu a jej vývoj neustále napreduje. Spolahlivé pripojenie dát je nevyhnutné pre čoraz inteligentnejšie, automatizovanejšie technológie. Mobilné dáta nám umožňujú rýchly prenos a v plne prepojenom digitálnom svete budeme potrebovať prepojiť doslova všetko, vrátane ľudí cez vozidlá, senzory, dáta na úložiskách. Piata generácia bezdrôtových sietí, ktoré sa momentálne vyvíjajú a sú nasadzované, ponúkajú výrazný pokrok ďalej ako je LTE. Avšak tieto siete nemusia byť schopné pokryť požiadavky na úplne prepojenie, ktoré bude nutné pre našu digitálnu spoločnosť [1].

Preto sa teda úsilie a snaha z priemyslu a akademickej oblasti zamerala na hľadanie niečoho, čo bude viac ako 5G a tak vznikajú prvé koncepty 6G siete. Je tu vysoký predpoklad, že práve 6G siete podstúpia akúsi transformáciu, pomocou ktorej sa budú jednoznačne líšiť od predchádzajúcich bezdrôtových generácií celulárnych systémov [2], [3]. Na rozdiel od predchádzajúcich generácií bude 6G transformačná a spôsobí revolúciu v bezdrôtovom vývoji. Tu sa posunieme od pripojených vecí k pripojenej inteligencii, ktorá bude mať oveľa prísnejšie požiadavky. Medzi ne patria veľmi vysoké dátové rýchlosti až do 1Tb/s v dôsledku veľkého počtu zariadení. Ďalej tam bude potrebné mať veľmi vysokú energetickú účinnosť so schopnosťou podporovať zariadenia IoT bez batérií, ako aj masívnu kontrolu s nízkou odozvou (menej ako 1ms end-to-end) pre dosiahnutie vysokej rýchlosti. T akisto je dôležité dosiahnuť veľmi široké frekvenčné pásma (napríklad 76 GHz - 140GHz a 1THz - 3THz) na zabezpečenie vysokých dátových rýchlostí. Ak chceme vytvoriť širokopásmovú bezdrôtovú sieť, tak bude potrebné využiť nielen pozemnú sieť, ale takisto aj bezdrôtové pripojenie so satelitnými systémami. Poslednou nie menej podstatnou požiadavkou je prepojenie inteligencie so schopnosťou strojového učenia (Machine learning).

Očakáva sa, že do roku 2030 sa naša spoločnosť posunie smerom k oveľa viac digitalizovanej, dátovo riadenej a inteligentne inšpirovanej spoločnosti, ktorá bude potrebovať takmer okamžité a všadeprítomné bezdrôtové pripojenie na Internet. V najbližších rokoch teda vznikne niekoľko nových aplikácií, ktoré poskytnú takúto integráciu. Medzi ne by mali patriť napríklad virtuálna a rozšírená realita, 8K video streaming, hologramy, operácie na diaľku, priemysel 4.0, inteligentné domy, fog a edge Computing, integrované služby umelej inteligencie alebo bezpilotné lietadlá (Unmanned Aerial Vehicles – UAV), či napríklad také autonómne vozidlá [2]. Zároveň sa počíta s využitím mobilných viacpreskokových sietí

(MANET- Mobile Ad-Hoc Network, DTN - Delay Tolerant Network, VANET - Vehicular Ad-Hoc network, FLYNET a DRONET) ako s progresívnymi sieťami na prenos dát medzi jednotlivými terminálmi.

Viacpreskokové mobilné siete boli navrhnuté pre špecifické typy aplikácií a špecifické oblasti, kam bežné siete s infraštruktúrou nedokážu poskytovať svoje služby. Pre potreby 6G sietí sa budú využívať aj D2D (Device-to-Device) a VANET siete, ktoré majú za úlohu poskytnúť dátové služby aj do vzdialených oblastí, kam nedosahujú poskytovatelia komerčných služieb. Sieť D2D [4] umožňuje bezdrôtovú komunikáciu medzi samotnými senzormi v Internete vecí (IoT) bez potreby sietí s infraštruktúrou a VANET siete umožňujú komunikáciu medzi pohybujúcimi sa zariadeniami. Špecifickým typom sietí sú DTN siete, ktoré umožňujú prenos dát aj v prípade sporadickkej konektivity medzi jednotlivými mobilnými terminálmi. MANET uzly (terminály) vyžadujú spoločné zdieľanie sieťových zdrojov a pre spoľahlivý a bezpečný prenos dát. Nesprávne fungujúce uzly by však mohli narušiť bezpečnosť viacpreskovej siete tým, že budú konať sebecky a zlomyseľne, využívajúc spoluprácu iných uzlov a ich zdrojov [5].

V súčasnej dobe sa do popredia dostáva otázka využitia technológie Blockchain nielen pre siete 6. generácie, ale aj pre rozličné typy MANET sietí. Blockchain poskytuje nové mechanizmy na eliminovanie zlomyseľných mobilných uzlov a na zvýšenie dôveryhodnosti samotnej dátovej komunikácie bez potreby dôveryhodných tretích strán. Tieto vlastnosti by mohli umožniť spolupracujúcim uzlom pripojiť sa a opustiť MANET sieť bez predchádzajúcich informácií o ich dôveryhodnosti a pritom im poskytnúť určitú formu odmeny za spoluprácu.

#### A. Výzvy pre siete 6. generácie

Medzi najdôležitejšie výzvy pri budovaní 6G siete patria jednoznačne nasledujúce problémy, s ktorými sa treba vysporiadať. Medzi ne patria [2]:

##### **Masívne pripojenie v budúcich systémoch**

*Výpočtovo orientovaná komunikácia* je zásadnou požiadavkou pre budúce ekosystémy. Komunikácia medzi zariadeniami ako aj medzi strojmi si vyžaduje presnosť s takmer nulovým oneskorením pre precíznosť svojich operácií. *Synchronizácia* je významnou požiadavkou v časovo kritických priemyselných aplikáciách. Kritické chrbticové systémy v krajinách, vrátane systémov na distribúciu energie a automobilových sietí sa musia synchronizovať v reálnom čase pre to, aby ich prevádzka bola presná. *Vyššia priepustnosť* je dôležitá pre kritické systémy, ktoré budú využívať komunikačné ekosystémy 5G a novšie si vyžadujú súčasné prepojenie veľmi veľkého počtu zariadení. To znamená, že sieťová infraštruktúra ako taká by mala zvládnuť obrovský objem transakcií v reálnom čase.

##### **Bezpečnostné požiadavky v budúcich ekosystémoch**

*Dôvernosť* je veľmi dôležitý nástroj pre budúcu infraštruktúru ako je napríklad IoT, ktorú vystavuje bezdrôtové pripojenie obrovským hrozbám. Šifrovacie algoritmy musia byť jednoduché pre zariadenia s nízkou spotrebou IoT. *Integrita* je nutná pri produkovani veľkého množstva údajov kvôli tomu, aby k týmto údajom mali prístup oprávnení používatelia, ktorí ich budú môcť upravovať pri ich ďalšom prenose. *Dostupnosť* služieb je základnou požiadavkou pre budúce siete. Najmä sofistikovanosť 5G ekosystémov s veľkým objemom vzájomne prepojených zariadení rozširujú riziko útokov DDoS. *Autentifikácia a kontrola prístupu* je dôležitá pre údaje, ktorá sa prenášajú alebo ukladajú. Tu je potrebné zabezpečiť pomocou mechanizmov kontroly a prístupov, aby sme zabránili neoprávnenej manipulácii. *Audit* sa vyžaduje na vyhodnotenie správania sa nájomníkov v sieťovom ekosystéme, ktorí ho využívajú. V prípade zvýšených bezpečnostných štandardov môže byť na identifikáciu a označenie správania týchto nájomcov potrebný audit na úrovni paketov.

##### **Vyššia spotreba dát v sofistikovaných riešeniach**

Vyššia prenosová rýchlosť je jedným z najvýznamnejších očakávaní v budúcich sieťových ekosystémoch. Aplikácie ako VR, holografická komunikácia, 16K video a 3D ultra video vyžadujú vyššiu rýchlosť a spotrebu dát.

##### **Obmedzenia prostriedkov zariadenia**

Očakáva sa, že výpočtové a úložné obmedzenia obmedzia možnosti kryptografických algoritmov a nakoniec povedú k odchýlke od štandardných mechanizmov. Štandardné prijatie zabezpečenia je pri takýchto obmedzeniach prostriedkov zariadenia ťažšie.

## Blockchain pre siete 6G

## Výzvy v oblasti 6G

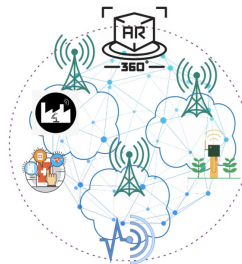
- Masívne konektivita systémov
- Požiadavky na bezpečnosť so škálovateľnosťou
- Vysoká spotreba dát budúcich prenajímateľov
- Obmedzenia zdrojov zariadení
- Požiadavky na prevádzkyschopnosť a integrovanie

## Výhody blockchainu pre 6G

- Inteligentná správa zdrojov
- Zvýšené bezpečnostné funkcie
- Škálovateľnosť

## Aplikácia blockchainu v 6G

- Priemyselné použitie v Priemysle 4.0
- Bezproblémové monitorovanie a ochrana prostredia
- Inteligentná zdravotná starostlivosť
- Decentralizovaná a dôveryhodná 6G komunikačná infraštruktúra a riešenia



## Možnosti výskumu v budúcnosti

- Internet všetkého
- Ukladanie a analýza údajov
- Umelá inteligencia
- Komunikácia medzi vozidlami (vehicle-to-vehicle)
- Bezpečné lietadlá

Obr. 1 Úloha Blockchainu v 6G sieťach.

## II. TECHNOLÓGIA BLOCKCHAIN PRE VIACPRESKOKOVÉ MOBILNÉ SIETE V 6G SIETI

Prvá generácia digitálnej revolúcie dala možnosť vzniknúť internetu informácií a dostala ho do popredia. Druhá generácia bude poháňaná decentralizovanou technológiou blockchainu, prináša pre nás hodnotný internet - skutočnú platformu typu peer-to-peer. Táto platforma má vysoký potenciál ísť ďaleko za hranice digitálnej meny a zaznamenať pre nás ľudí distribuované prakticky všetko hodnotné bez silných sprostredkovateľov. Niektorí dokonca hovoria o decentralizovanej technológii blockchain ako o alchýmii dvadsiateho prvého storočia, pretože môže využívať vybavenie koncového používateľa na prevod výpočtovej techniky do digitálneho zlata. Blockchain sa preto považuje za nevyhnutnú technológiu na vybudovanie dôvery pre budúce siete. Pretože sa blockchain považuje za jednu z kľúčových podporných technológií pre mobilné siete 6G, je takisto nevyhnutné preskúmať rôzne výhody, príležitosti a výzvy, ktoré sa pri jeho využívaní očakávajú. Na Obr. 1 je zobrazená úloha Blockchainu v 6G sieťach [2].

Úloha blockchainu v 6G sieťach sa týka priebežného vytvárania vertikálnych služieb, diferencovaných služieb s vytváraním sietí. Jednou z dôležitých úloh blockchainu a distribovanej knihy (Distributed ledger technology – DLT) ako novej generácie služieb distribovaného snímania pre aplikácie riadenia 6G, ktorých potreba pripojenia si bude vyžadovať kombináciu ultra spoľahlivej komunikácie s nízkou latenciou (Ultra Reliable Low Latency Communications – URLLC) a masívnej komunikácie strojového typu (massive Machine Type Communications – mMTC), ktoré majú zabezpečiť spoľahlivé pripojenie, škálovateľnosť a hlavne nízku odozvu [3]. Na Obr. 1 je zobrazená úloha Blockchainu v 6G sieťach.

Blockchain je jednou z najvýraznejších technológií, ktorá má na starosti ukázať plný potenciál systémov v 6G sieťach. Schopnosti a silné stránky technológie slúžia na elimináciu potencionálnych výziev. Medzi ne patria napríklad [2], [4]:

- **Správa spektra** - zavedená dynamická politika správy licencií na frekvenčné spektrum s cieľom efektívne riadiť dané frekvenčné spektrum. Umožňuje sekundárnym používateľom bez licencie oportunistický prístup k licencovanému spektru bez zasahovania do licencovaného primárneho používateľa.
- **Inteligentné riadenie zdrojov** - sa zaoberá správou sieťových zdrojov, ktorá je náročná z hľadiska očakávaných veľmi veľkých požiadaviek na pripojenie pre budúce telekomunikačné ekosystémy. Zdieľanie spektra a decentralizovaný výpočet musia byť kompatibilné s očakávanou masívnou infraštruktúrou.
- **Zvýšenie sieťovej a dátovej bezpečnosti** - venuje sa problematike ochrany osobných údajov, overovaní a kontrola prístupu, integrity dát obrovského obrovského objemu, využívanie technológie Blockchain ako prevencie útokov DDoS a pod..
- **Škálovateľnosť** - Limitácia škálovateľnosti centralizovaných systémov môže byť obmedzená použitím blockchainu a smart kontraktov, aby mohli čeliť očakávanému veľkému dopytu po pripojení v budúcnosti.

### A. Decentralizácia prostredníctvom blockchainu

Rastie predovšetkým záujem o prispôsobenie blockchainu pre konkrétne potreby IoT s cieľom vyvinúť rôzne druhy aplikácií Internetu vecí, ktoré budú založené na technológií využívajúcej práve blockchain ako taký (Blockchain-based Internet of Things). Ako príklad môžeme uviesť inteligentné mestá, priemysel 4.0 (Industry 4.0) alebo finančné transakcie a poľnohospodárstvo [4]. Niektorí autori poukázali na veľmi dôležitú úlohu inteligentných zmlúv (smart contracts), ktoré boli definované ako časti sebestačného kódu, ktorý sa vykonáva autonómne pri splnení určitých podmienok. Ethereum sa javilo ako najpopulárnejšia platforma na vykonávanie inteligentných zmlúv, ktorá využíva túto technológiu.

Používanie Etherea umožní používateľom písať a spustiť svoj vlastný kód na vrchnej úrovni siete. Aktualizáciou kódu dokážu používatelia upraviť správanie zariadení IoT kvôli zjednodušenej údržbe a oprave chýb. Okrem známych problémov BIoT, ako je napríklad hostovanie blockchainu na zariadeniach IoT s obmedzenými zdrojmi, nízke transakčné rýchlosti a zdĺhavé vytváranie blokov sa identifikovali niektoré výzvy, ktoré presahujú počiatočný vývoj BIoT a ich nasadenie, bude ešte predmetom skúmania v blízkej budúcnosti. Jeden z najväčších problémov pre uvoľnenie potenciálu pre BIoT je decentralizované vlastníctvo.

V poslednej dobe sa niektoré štúdie začali hlbšie venovať niektorým nedostatkom BIoT. Zariadenia IoT s obmedzenými prostriedkami sa postupne odbremenili z výpočtovo náročných úloh ako sú napríklad vytváranie, pripájanie alebo sledovanie transakcií blockchainu pre výkonnejšie okrajové zdroje výpočtovej techniky (Edge Computing Resources) ako sú cloudlety (cloudlets). Ďalším problémom, nad ktorým sa autori zaoberali, je problém kritického riadenia prístupu. Tento problém spočíva v tom, aby sa zabránilo prístupu k prostriedkom BIoT pre neoprávnené subjekty. Preto sa využila platforma Ethereum na dosiahnutie rôznych metód pre riadenie prístupu. Konkrétne sa brány využili tak, aby fungovali ako BIoT agenti služieb (BIoT service agents) pre ich klaster miestnych zariadení IoT s obmedzenými prostriedkami a to tak, že ukladali svoje účty blockchain a používali ich na vykonávanie inteligentných zmlúv v ich mene. Navrhovaný rámec založený na týchto inteligentných zmluvách pozostáva z viacerých zmlúv o kontrole prístupu (Access Control Contracts – ACCs) [3], [4].

Každý ACC udržiava zoznam nesprávneho správania pre každý zdroj BIoT, vrátane podrobností a času nesprávneho správania, ako aj trestu na jeho tému - napr. blokovanie žiadostí o prístup na určité časové obdobie. Okrem toho ďalej zahŕňa sudcovskú zmluvu (Judge Contract – JC), ktorá zavádza určitú metódu posudzovania nesprávneho správania. Po prijatí správ o takomto chovaní od ACC, následne určí JC trest zodpovedajúcim subjektom a vráti rozhodnutia ACC na ich vykonanie.

Mnoho ďalších štúdií považovalo práve Ethereum za správnu platformu na využitie blockchainu. V predbežnej štúdií s využitím inteligentnej veci služby prenájmu ako príklad služby BIoT, autori porovnali nasledujúce 4 rôzne architektonické štýly založené na Ethereu.

Medzi ne patria:

- *Plne centralizovaný cloud bez blockchainu*
- *Pseudo-distribúované veci fyzicky uložené v centrálnom cloude*
- *Distribúované veci priamo ovládané inteligentnou zmluvou*
- *Plne distribuované systémy*

Predbežné výsledky naznačujú, že práve plne distribuovaná architektúra, v ktorej je koncový bod blockchainu nasadený na zariadení koncového používateľa, je vynikajúca z hľadiska robustnosti a bezpečnosti. Aby sa zabránilo riziku, že dôjde k odcudzeniu alebo napadnutiu tajných kľúčov na podpis, autor navrhol použiť zabezpečené prvky karty java a kryptomenovú inteligentnú kartu (Crypto Currency Smart Card – CCSC) [3], [4].

### B. Ethereum vs. Bitcoin

Typická blockchainová sieť je v podstate distribuovaná databáza obsahujúca záznamy o všetkých transakciách alebo digitálnych udalostiach, ktoré boli vykonané zúčastnenými stranami alebo boli medzi nimi zdieľané [3].

Blockchainy možno rozdeliť na verejné a súkromné siete. V prvej kategórii sa do blockchainu môže pripojiť a zúčastniť ktokoľvek. Naopak, súkromný blockchain uplatňuje určité mechanizmy kontroly prístupu, aby určil, kto sa môže do siete pripojiť. Verejný blockchain je nemenný, pretože žiadnu z transakcií nie je možné manipulovať alebo zmeniť. Takisto je pseudo anonymný, pretože identita účastníkov transakcie je reprezentovaná adresným kľúčom v podobe náhodného refazca [4] (Tabuľka 1).

Bitcoinový blockchain je určený prevažne na ulahčenie bitcoinových transakcií. Je to prvá plne funkčná digitálna mena na svete, ktorá je skutočne decentralizovaná, má otvorený zdrojový kód a je odolná voči cenzúre. Bitcoin využíva kryptografický mechanizmus konsenzu

	Verejný blockchain	Súkromný blockchain
Typ siete	Plne centralizovaná	Čiastočne centralizovaná
Prístup	Bez oprávnenia na čítanie/zápis	S oprávnením na čítanie/zápis
Identita používateľa	Pseudoanonymný	Známí účastníci
Konsenzuálny mechanizmus	Dôkaz práce/ dôkaz podielu	Vopred schválení účastníci
Určenie konsenzu	Všetky zariadenia potvrdzujúce platby	Jednou organizáciou
Nemennosť	Takmer nemožné manipulovať	Môže byť manipulované
Využitie	Akékoľvek decentralizované aplikácie	Obchodné aplikácie

Tabuľka 1

Porovnanie verejného a súkromného blockchainu

s dôkazom o práci (Proof-of-work – PoW), ktorý je založený na hashovacej funkcii SHA-256 a digitálnych podpisoch. Dosiahnutie konsenzu poskytuje extrémnu úroveň odolnosti voči poruchám, zabezpečuje nulový výpadok a spôsobuje, že údaje uložené v blockchaine sú navždy nemenné a odolné voči cenzúre, pretože každý môže vidieť históriu blockchainu vrátane akýchkoľvek údajov alebo správ. Existujú dva rôzne typy aktérov, ktorých úlohy sú definované nasledovne [4].

- *Pravidelné uzly (Regular nodes)*
- *Uzol používaný na potvrdenie platby (Miners)*

*Pravidelný uzol* je bežný aktér, ktorý má len kópiu blockchain a používa sieť blockchain na odosielanie alebo prijímanie bitcoinov. *Uzol používaný na potvrdenie platby* je aktér s konkrétnou úlohou, ktorý vytvára blockchain prostredníctvom overovania transakcií vytváraním blokov a ich odosielaním do siete blockchain, aby boli zaradené ako bloky. Ťažiar slúžia ako ochrancovia siete a môžu pôsobiť z ktoréhokolvek miesta na svete, pokiaľ majú dostatočné znalosti o procese ťažby, hardvéri a softvéri potrebné na túto činnosť a internetové pripojenie.

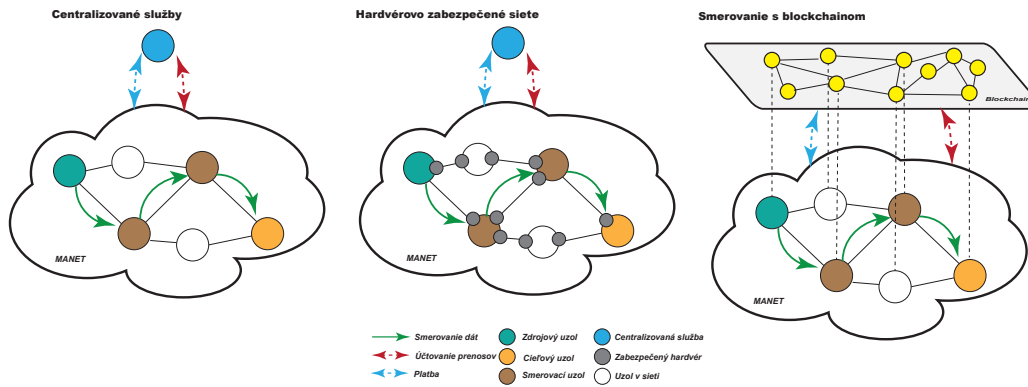
V blockchaine Bitcoinu sa blok ťaží približne každých 10 minút a veľkosť bloku je obmedzená na 1 MByte. Blockchain Bitcoinu je obmedzený na rýchlosť siedmich transakcií za sekundu, čo ho robí nevhodným pre vysokofrekvenčné obchodovanie. Medzi ďalšie nedostatky bitcoinového blockchainu patrí jeho skriptovací jazyk, ktorý ponúka len obmedzený počet malých inštrukcií a je neúplný. Okrem toho vývoj aplikácií využívajúcich Bitcoin skriptovacieho jazyka vyžaduje pokročilú zručnosť v oblasti programovania a kryptografie. Ethereum je v súčasnosti po Bitcoine druhý najpopulárnejší verejný blockchain. Ethereum má svoju vlastnú kryptomenu s názvom Ether, ktorá poskytuje primárnu formu likvidity, umožňujúcu výmenu hodnoty v rámci celej siete. Ether tiež poskytuje mechanizmus na platenie a získavanie transakčných poplatkov, ktoré vznikajú pri podpore a používaní siete.

Podobne ako Bitcoin, aj Ether bol predmetom špekulácií, ktoré boli svedkami veľkých výkyvov. Ethereum je vhodné na vývoj decentralizovaných aplikácií (Decentralized applications – DApps), ktoré je potrebné rýchlo vytvoriť, efektívne a bezpečne s nimi komunikovať prostredníctvom platformy blockchain. Podobne ako Bitcoin, aj Ethereum používa metódu konsenzu PoW na overovanie transakcií a dokazovanie dosiahnutia určitého množstva práce. Hashovací algoritmus používaný mechanizmom PoW sa nazýva Ethash [4].

### C. Blockchain technológia v procese smerovania v MANET sieti

V MANET sú smerovacie protokoly zodpovedné za proces objavenia, udržiavania komunikačnej cesty a za prenos dát medzi zdrojovým a cieľovým uzlom. Smerovacie protokoly musia poskytovať QoS služby, ochraňovať sieť pred preťažením smerovacími paketmi a zároveň musia vykonať procesy na potvrdenie totožnosti jednotlivých uzlov. Bezpečnosť je preto nevyhnutným aspektom, ktorý smerovací protokol neposkytuje, a je potrebné ju primerane pridať. Existujúce bezpečnostné algoritmy a protokoly si vyžadujú výmenu a ukladanie ďalších informácií, ktoré môžu vyčerpať samotné zdroje jednotlivých uzlov.

Blockchain teda možno definovať ako decentralizovaný proces vynájdený s blokmi inteligentného kódu a určitým stupňom šifrovania [5]. Blockchain umožní poskytovanie nových typov služieb, ktoré poskytnú silné zabezpečenie nielen smerovacích paketov, ale aj smerovacích dát, identifikáciu zlomyselných a nebezpečných uzlov, interoperabilitu. Smerovacie protokoly s implementovanou technológiou blockchain umožnia vytvorenie aktuálnych komunikačných ciest na základe aktuálnych transakcií a zároveň umožní autentifikáciu a identifikáciu samotných mobilných uzlov využitím hashovacích blokov, uložených v smerovacích paketoch (Obr. 2).



Obr. 2 Využitie technológie blockchain v MANET.

Výskum využitia blockchainu pre MANET siete je len v počiatku. Cieľom výskumných aktivít je poskytnúť bezpečné a spoľahlivé prostredie pre rozličné typy mobilných viacpreskokových sietí s dôrazom kladeným na mobilitu samotných uzlov. Samotná technológia blockchain nám umožní riešiť problematiku autentifikácie prepojenia jednotlivých typov mobilných viacpreskokových sietí tak, aby sa vytvorila decentralizovaná dynamická sieť [6], [7], [8]. Táto sieť si bude uchovávať aktuálne informácie o aktuálnom stave všetkých blokov v danom reťazci a zároveň bude zvyšovať dôveryhodnosť jednotlivých mobilných uzlov a sieťovú bezpečnosť.

### III. ZÁVER

Tento článok sa venuje problematike využitia technológie blockchain pre siete 6. generácie, ktorý sa v súčasnosti vyvíja pre mobilné a bezdrôtové komunikačné technológie. Pozornosť sa upriamila na konkrétne problémy, ako sú nové požiadavky a jej výzvy, ktorými sa bude treba zaoberať pri tejto technológii. Jednou z najdôležitejších výziev je zvýšenie sieťovej a dátovej bezpečnosti v 6G sieťach. Technológia Blockchain je považovaná za jednu z najprogressívnejších sa vyvíjajúcich sa technológií, pomocou ktorej sa ukáže jej plný potenciál práve pre siete 6. generácie. Významnú úlohu zohráva aj jej využitie v procese smerovania a prenosu IoT dát prostredníctvom viacpreskokových mobilných sietí (MANET, DTN, DRONET, FLYNET a pod.). Tento článok sa venuje základnej myšlienke, ktorej sa budeme naďalej venovať v našej vedeckej činnosti.

### POĎAKOVANIE

Táto publikácia vznikla vďaka podpore Agentúry na podporu výskumu a vývoja v rámci projektu č. APVV-17-0208 - Odolné mobilné siete na doručovanie obsahu.

### LITERATÚRA

- [1] WANG, Minghao, et al. Security and privacy in 6G networks: New areas and new challenges. *Digital Communications and Networks*, 2020, 6.3: 281-291.
- [2] HEWA, Tharaka, et al. The role of blockchain in 6G: challenges, opportunities and research directions. In: *2020 2nd 6G Wireless Summit (6G SUMMIT)*. IEEE, 2020. p. 1-5.
- [3] XU, Hao, et al. Blockchain-enabled resource management and sharing for 6G communications. *Digital Communications and Networks*, 2020, 6.3: 261-269.
- [4] MAIER, Martin; EBRAHIMZADEH, Amin. *Toward 6G: A New Era of Convergence*. John Wiley and Sons, 2021.
- [5] LAUBE, A., et. al. A solution to the split and merge problem for blockchain-based applications in ad hoc networks, *International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks*, pp. 1-6, 2019.
- [6] LeMAHIEU, C.: *A Feeless Distributed Cryptocurrency Network*, 2018.
- [7] CORDOBA, D., et al. Blockgraph: A blockchain for mobile ad hoc networks, *2020 4th Cyber Security in Networking Conference (CSNet)*, 2020, pp. 1-8, doi: 10.1109/CSNet50428.2020.9265532.
- [8] HE, Y, et al. Blockchain based Truthful Incentive Mechanism for Distributed P2P Applications. *IEEE Access* 2018(6), 27324 – 27335



# Web kiosk v OS Linux

Filip Mondok, Anton Baláž, Emília Pietriková

Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

*filip.mondok@student.tuke.sk, anton.balaz@tuke.sk, emilia.pietrikova@tuke.sk*

**Abstrakt** — Zámernom tejto práce je vytvoriť systém na bezpečné spúšťanie mobilného kódu v prostredí web prehliadača. Na základe analýzy súčasného stavu problematiky a podobných web kiosk systémov sme vytvorili návrh prenosného kódu, ktorý zmení nastavenia počítača a vytvorí z neho web kiosk zariadenie. V prvej časti je možné nájsť jednotlivé úlohy, ktoré bolo treba spracovať z pohľadu navrhovaného riešenia. Druhá kapitola sa zameriava na analýzu už existujúcich riešení a analýzu možností, ktoré by sme mohli použiť pri návrhu mobilného kódu. Tretia kapitola je zameraná na vyhodnotenie analýzy, vďaka čomu sme dokázali vytvoriť návrh a výsledné riešenie prenosného systému.

**Kľúčové slová** — Ubuntu, Kiosk, Web, Linux, GNOME

## I. ÚVOD

Pri tvorbe riešenia sme sa zamerali na vytvorenie web kiosk systému, ktorý je prenosný medzi viacerými distribúciami Linux operačného systému. Vďaka týmto zariadeniam je možné zmodernizovať buď školské priestory alebo podniky a tak aj pomôcť zákazníkom alebo študentom pri získavaní informácií alebo objednávaní a rezervácií. Výsledkom práce je web kiosk prenosný kód, ktorý po spustení zablokuje určité funkcie operačného systému, ako napríklad klávesové skratky, alebo voľnú manipuláciu so systémom a otvorí webový prehliadač so stránkou vopred určenou vlastníkom. Mobilný kód bude spustiteľný v operačnom systéme Linux.

## II. TVORBA NAVRHOVANÉHO RIEŠENIA

### A. Analýza existujúcich riešení

Prvou z častí bolo oboznámenie sa s existujúcimi rovnakými alebo podobnými riešeniami. Pri týchto riešeniach sme zanalyzovali zložitejšie aplikácie, ktoré sa primárne využívajú v školách a umožňujú zablokovať zariadenie na diaľku no rovnako aj na jednoduché riešenia pre podnikateľov, ktorí si chcú vytvoriť vlastný web kiosk systém. Riešenia ktoré boli postavené na Linux operačnom systéme boli upravené distribúcie Linuxu a bolo ich potrebné bootovať do zariadenia.

Po úspešnej analýze a odsledovaní prvkov už vytvorených systémov sme dokázali odsledovať aké funkcie sa uzamykajú a ako fungujú. Vďaka čomu sme dokázali pokračovať v navrhovanom riešení práce a analyzovať jednotlivé časti systému a jeho aplikácie, ktoré budeme využívať.

### B. Analýza technológií

Po analýze existujúcich riešení sme analyzovali jednotlivé implementácie Linux distribúcií pri ktorých sme sa rozhodovali či náš systém bude vytvorený pre GNOME alebo KDE grafické prostredie. Každé z týchto prostredí malo svoje výhody ale aj nevýhody. Zároveň sú rôzne distribúcie ktoré využívajú GNOME alebo KDE prostredie, kde operačný systém Ubuntu jeden z najznámejších distribúcií Linuxu využíva práve GNOME.

Postupne sme analyzovali Mandatory access control, k čomu slúži aké má pre našu prácu využitie a aké aplikácie s nim pracujú. Využíva sa primárne pre riadenie prístupu používateľov a funkcií k súborom alebo aplikáciám v systéme. Vďaka tomu je možné zabezpečiť systém pred nechceným používaním od používateľov [5]. Mandatory access control je možné vytvoriť a spravovať za pomocou viacerých aplikácií ako napríklad SeLinux, Apparmor [3] alebo systém Tomoyo.

Ďalším z analyzovaných častí bol sandboxing vďaka ktorému je možné bez narušenia systému spúšťať aj škodlivé kódy, ktoré za iných okolností v systéme by mohli poškodiť systém. No podstatou samotného pieskoviska je vytvorenie virtuálneho priestoru, z ktorého nie je možné ovplyvniť systém tak vďaka tomu je možné systém ochrániť. Využíva sa hlavne pri spúšťaní neoverených kódov, aplikácií alebo webových stránok [7].

Analyzovali sme rovnako aj webové prehliadače Google Chrome alebo Mozilla Firefox kde sme zisťovali, ktorý z nich bude pre web kiosk bude najvhodnejší a najbezpečnejší. No rovnako aj možnosti filtrovania webových stránok, aby sme povolili iba stránky, ktoré administrátor bude vyžadovať. Túto možnosť nám poskytovali proxy servery ako napríklad Privoxy, Squid alebo Tinyproxy. Pomocou týchto proxy serverov sme dokázali vytvoriť blacklist, čo nám umožnilo vytvoriť zoznam so stránkami, ktoré nemajú byť prístupné alebo whitelist ktorý mal opačnú funkcionálnosť a to povolilo stránky, ktoré boli v zozname a ostatné zablokovalo [4].

Ako posledné sme analyzovali možnosti potlačania funkcionálnosť klávesových skratiek alebo ich zablokovanie. Je to dôležitá časť, pretože v Linux operačnom systéme dokážeme pomocou klávesových skratiek otvoriť terminál, alebo iné aplikácie a tak by sa užívateľ mohol dostať mimo web kiosk systému. Analyzovali sme dve možnosti blokovania klávesových skratiek a to pomocou Dconf alebo Xmodmap.

### C. Vyhodnotenie Analýzy

Ako ďalší krok sme zvolili po analýze technológií vyhodnotenie tejto analýzy. Kde sme sa rozhodli, ktoré z operačných systémov a aplikácií použijeme. Pri vyhodnotení, ktorý z grafických prostredí operačného systému Linux sme sa rozhodli pre GNOME pre jeho časté používanie v školách a jeho stabilitu.

Webový prehliadač sme vybrali Google Chrome pre jeho kiosk mód a kompatibilitu so systémom Linux a ako proxy server na filtrovanie webstránok sme vybrali Tinyproxy, pretože nezaberá veľa miesta v zariadení a je jednoducho nastaviteľný.

Mandatory Access control a sanddox v našom systéme nebolo potrebné využívať, pretože používateľ web kiosk zariadenia nebude mať prístup k ničomu inému ako obmedzenému webovému prehliadaču.

Blokovanie klávesových skratiek sme sa rozhodli využiť Xmodmap, pri ktorom sme potrebovali zistiť jednotlivé kódy kláves ktoré chceme zablokovať. Keďže po skončení web kiosk systému má zariadenie fungovať ako pred spustením nášho systému tak je potrebné tieto klávesy naspäť nastaviť.

## III. NÁVRH SYSTÉMU

Ešte predtým ako sme začali vytvárať náš systém sme si ho museli navrhnuť, aké funkcie by mal poskytovať, jeho štruktúru a fungovanie. Ako prvou časťou sme si vytvorili diagram pomocou ktorého sme dokázali náš systém rozdeliť na jednotlivé časti.

- Proxy server
- Klávesové skratky
- Procesy bežiacie na pozadí
- Zrkadlenie monitorov
- Ukončenie kiosk módov

Proxy server sme potrebovali zahrnúť a spúšťať a aké funkcie využijeme pri jeho nastavovaní. Aby sme dokázali filtrovať webové stránky.

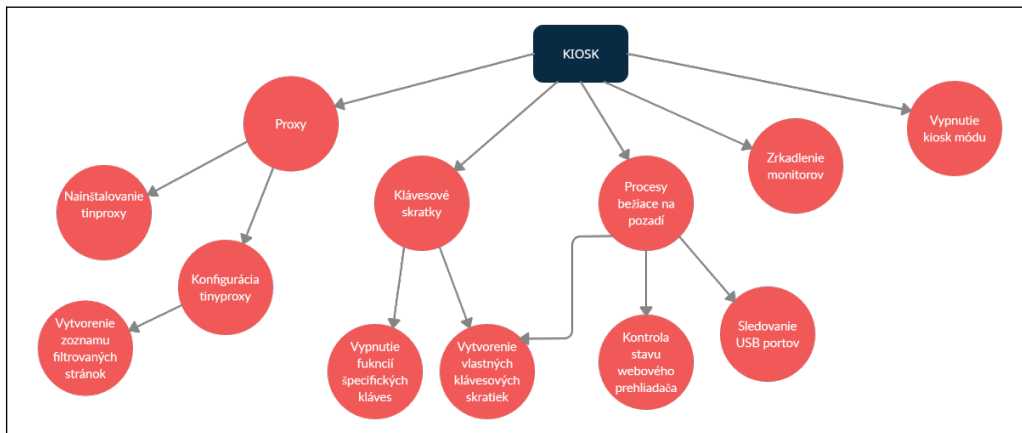
Vytvárali sme zoznam kláves ktoré budeme blokovať. Tieto klávesy sme získavali zo zoznamu klávesových skratiek Google Chrome, Gnome, Ubuntu a testovaním priamo v systéme [6]. No zároveň sme potrebovali vytvoriť aj vlastnú klávesovú skratku pomocou ktorej sme náš systém dokázali ovládať. Túto skratku sme navrhli, aby sa vytvorila pomocou jazyka Python a knižnice Pyinput.

Vo web kiosk systéme sme navrhli aj kontrolné procesy, ktoré bežia na pozadí počas behu systému. Medzi tieto procesy patrí kontrola stavu webového prehliadača či je vypnutý alebo zapnutý, pokiaľ sa vypne tak ho zapne znova, sledovanie USB portov či nebolo pridané nové zariadenie do systému, kde po pridaní zariadenia sa vykoná kód na zablokovanie kláves nanovo. Rovnako aj sledovanie či nebola stlačená vytvorená klávesová skratka.

Navrhli sme aj funkciu zrkadlenia monitorov je vhodná pokiaľ majiteľ kiosk zariadenia potrebuje mať viacero monitorov na jedno zariadenie, tak jednoducho ich dokáže odzrkadliť.

Pri vypínaní web kiosku systému je dôležité, aby zariadenie sa dostalo do pôvodného stavu, ako napríklad vrátenie kláves do pôvodného stavu alebo vypnutie proxy servera a všetkých procesov, ktoré bežali na pozadí počas spusteného systému.

Obrázok 1 Návrh štruktúry systému



#### IV. IMPLEMENTÁCIA SYSTÉMU

Cieľom implementácie a návrhu bolo vytvorenie prenosnej web kiosku aplikácie, ktorá umožní majiteľovi zariadenia s Linux operačným systémom zamedziť prístup užívateľom do systému a povoliť prístup iba na špecifické stránky vo webovom prehliadači. Z analýzy a návrhu sme vybrali GNOME grafické prostredie, na ktorom staviame našu aplikáciu. Ako sme už vyššie spomínali, grafické prostredie GNOME využíva Debian, Ubuntu a Mint operačné systémy.

V aplikácii sme sa rozhodli využívať shell skripty a Python jazyk, vďaka ktorým sme dokázali zabezpečiť našu aplikáciu pred nechceným používaním. Shell využívame, pretože je to jazyk, na ktorom stojí samotný Linux, a Python z dôvodu, že je to vysokoúrovňový jazyk, ktorý má k dispozícii mnoho knižníc pracujúcich so systémom Linux.

Aplikácia je zložená z viacerých častí, a to napríklad updatovanie systémových balíčkov, doinštalovanie potrebných knižníc, ktoré naša aplikácia vyžaduje, zablokovanie kláves, zapnutie webového prehliadača vo full screene alebo aj vytvorenie vlastných klávesových skratiek, pomocou ktorých vypneme web kiosku. Všetky tieto funkcie zabezpečujú, aby bola naša aplikácia funkčná a zabezpečila počítač pred zneužitím.

Štruktúra systému je rozdelená na viacero samostatných častí, kde každá časť vykonáva vlastnú úlohu, prípadne spúšťa druhú časť. Kde pri inštalácii web kiosku systému do zariadenia sa aktualizuje systém a nainštalujú sa potrebné balíčky pre beh aplikácie. Medzi potrebnými balíčkami sa nachádza aplikácia Google Chrome, a ďalšie Python knižnice, ktoré potrebujeme ku behu aplikácie. Rovnako aj inštalujeme aj Tinyproxy kde ho stiahneme z oficiálneho GIT repozitára, skompilujeme ho a nainštalujeme. Ďalej vytvoríme v systéme konfiguračný súbor do ktorého vložíme naše nastavenia Tinyproxy.

Po spustení systému ešte pred tým, ako začneme vykonávať prvý krok, sa musíme uistiť, že nie je zapnutý Tinyproxy. Po vypnutí Tinyproxy vypneme všetky procesy Google Chrome-u, pokiaľ bol zapnutý, vypneme klávesové skratky, ktoré by mohol užívateľ využiť pri zablokovaní ďalších procesov a spustíme proces, kde sa za pomoci Selenium drivera nakonfiguruje webový prehliadač a spustí sa v kiosku móde. Počkáme na spustenie a spracovanie doterajších procesov a zavoláme procesy pre kontrolu stavu webového prehliadača, sledovanie USB portov a zapnutie klávesových skratiek, pomocou ktorých je možné našu aplikáciu vypnúť.

##### A. Jednotlivé funkcie systému

Celý systém sa skladá z viacerých častí systému ktoré bolo potrebné implementovať pre jeho funkčnosť. Keďže využívame webový prehliadač bolo potrebné ho ako prvé nastaviť zapínanie do kiosku módu a rovnako aj kontrola či je počas behu systému aktívny pokiaľ nie je tak, aby sa zapol. Preto pred zapínaním systému vypíname všetky procesy od Chrome.

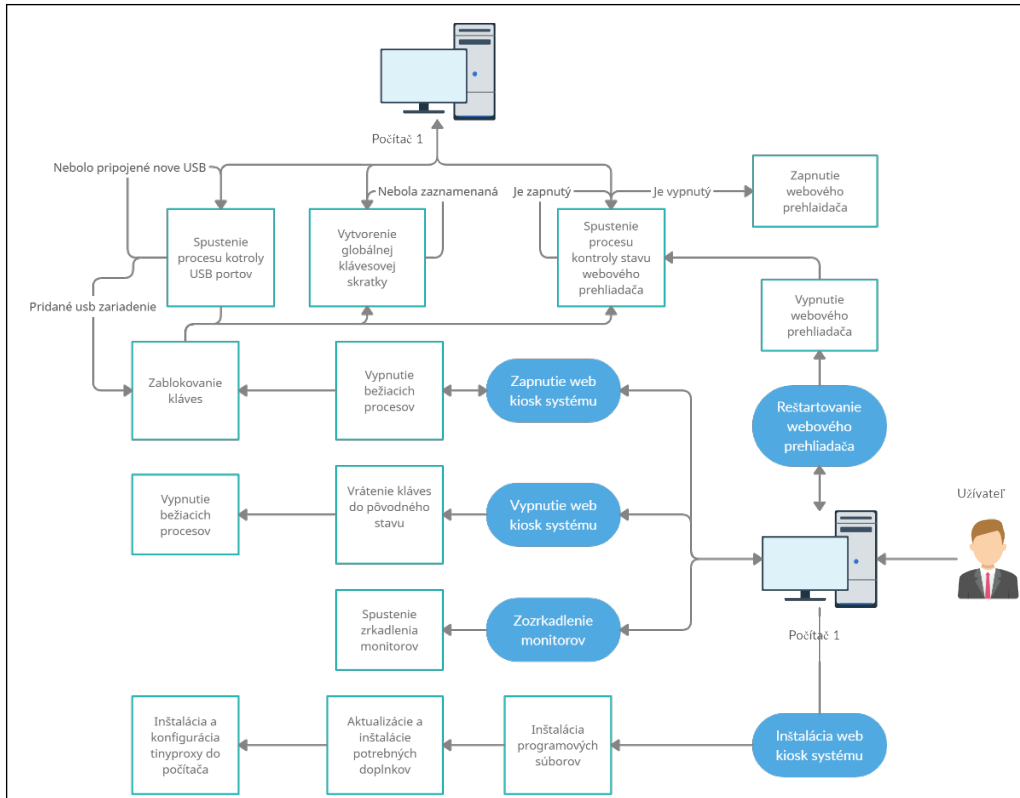
Ďalším krokom bola konfigurácia proxy, kde sme museli nastaviť jeho parametre, stránky ktoré má blokovať a aký druh blokovania požadujeme. Taktiež pri inštalácii sme potrebovali konfiguračné súbory, za pomoci ktorých sa proxy zapne.

Ako ďalší krok sme zablokovali klávesy, aby nebolo možné využívať funkcie klávesových skratiek a tak obísť systém náš systém a dostať sa ku iným aplikáciám alebo súborom ako je

webový prehliadač. Ako kontrolu sme pridali aj sledovanie USB portov kde po zaznamenaní že bolo pridané zariadenie, znova sa vykoná kód pre blokovanie kláves.

Klávesové skratky sme nie len blokovali, ale za pomoci Python knižnice sme implementovali funkcie ako vypnutie kiosk systému, reštartovanie webového prehliadača alebo odzrkadlenie monitorov. Po stlačení kombinácie kláves sa zobrazí menu s týmito funkciami a po zvolení funkcie je potrebné overiť prednastaveným heslom. Po overení sa funkcia vykoná.

Obrázok 2 Implementácia systému



## V. VYHODNOTENIE APLIKÁCIE DIGI ZŠ

Zhotovenú aplikáciu bolo potrebné preskúšať a vyhodnotiť. Z tohto dôvodu sa aplikácia nechala otestovať, aby sa preverila jej funkčnosť. Aplikáciu testovalo päť osôb, ktorým bol po otestovaní aplikácie zaslaný dotazník, kde bolo 5. otázok zameraných na praktickosť a funkčnosť aplikácie.

Otázky boli kladené zrozumiteľne, aby na otázky vedeli odpovedať rôznorodé vekové kategórie, ktorými boli účastníci tohto testu. Vyskytovali sa otázky zamerané na funkčnosť, či nedochádzalo k nečakaným chybám v systéme, alebo aj ako sa účastníkom pozdávala aplikácia a či bola vhodná pre kiosk zariadenia v školách alebo podnikoch. Na všetky otázky som dostal pozitívnu odpoveď.

Pri otázke, či sa im podarilo dostať mimo webového prehliadača, všetci odpovedali – NIE a rovnako aj na konfiguráciu systému pomocou návodu odpovedali, že sa im podarilo ho nakonfigurovať. Týmito odpoveďami sa môže celkový systém považovať za úspešný. Výsledkom tohto dotazníka bolo zistenie správnej funkčnosti systému a zároveň pohľad iných, nezávislých ľudí, ktorí tento systém testovali.

## VI. ZÁVER

V tejto práci sme sa zaoberali vytvorením prenosného kódu, ktorý po spustení premení zariadenie s Linux operačným systémom na web kiosk zariadenie. Rovnako sme analyzovali možnosti implementácie web kiosk systému, kde sme vyberali vhodné grafické prostredie operačného systému Linux, webový prehliadač alebo aj možnosti filtrovania prístupu na webstránky. Rovnako sme analyzovali možné využitie sandboxingu v našom systéme. Ďalším čiastkovým krokom bolo vyhodnotenie analýzy, ktoré z analyzovaných technológií sú najvhodnejšie použiteľné pre naše riešenie. Následne, ako sa nám podarilo vyhodnotiť analýzu, začali sme vytvárať návrh nášho systému, kde sme museli dôkladne vybrať klávesové skratky, ktoré zablokujeme, navrhnuť funkcie, ktoré zabezpečia kiosk mód z vnútra a znemožnia zrušiť webový prehliadač, alebo dostať sa do systému a spúšťať funkcie, poprípade zamedziť prístup

na nepovolené stránky. V práci sme využili teoretické poznatky z analýzy a návrhu riešenia pre implementáciu prenosného kódu, ktorý vytvoril z počítača s Linux operačným systémom web kiosk zariadenie, kde tento systém musí bežať na jednej z troch distribúcií, a to Ubuntu, Debian alebo Mint. Pri vytváraní nášho systému sme využili Bash a Python3, vďaka ktorým sa nám podarilo uskutočniť implementovanie funkcií do systému. Medzi tieto funkcie patrí vypnutie funkcie kláves, vytvorenie filtrovania stránok, vytvorenie vlastných klávesových skratiek, sledovanie pripojenia nových USB portov alebo aj sledovanie stavu webového prehliadača. Ako posledný krok sme vytvorili Debian balíček zo spomínaných funkcií pre jednoduchšiu inštaláciu do systému. V poslednej časti práce sme sa venovali testovaniu nášho vytvoreného systému, kde sme zisťovali nedostatky systému alebo možné budúce implementácie. Test prebiehal s testovacími subjektami, všetci mali už skúsenosti s Linux operačným systémom. Subjekty boli so systémom spokojné, na samotný kiosk systém nemali žiadne námietky. Navrhli dve možné budúce implementácie do nášho systému. Jeden z nápadov sa týkal nastavovania kiosk módu a toho, že by bolo vhodné, aby systém mal len jeden konfiguračný súbor, a z technickej časti, kde by prijali to, ak by bolo možné zablokovať USB porty v zariadení. Medzi prínosy práce je možné zaradiť vytvorenie fungujúceho web kiosk módu v Linux operačnom systéme, ktorý je prenosný a jednoducho implementovateľný.

#### POĎAKOVANIE

Tento výskum bol podporený projektom 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market.

#### LITERATÚRA

- [1] Bc. Filip Mondok, Web kiosk v OS Linux, Diplomová práca, 24.4.2021
- [2] VAKULENKO, Yulia; OGHASI, Pejvak; HELLSTRÖM, Daniel. Innovative framework for self-service kiosks: Integrating customer value knowledge. *Journal of Innovation & Knowledge*. 2019, roč. 4, č. 4, s. 262–268.
- [3] NAKAMURA, Yuichi. SELinux & AppArmor-Comparison of Secure OSes. Hitachi Software Engineering. 2007. Dostupné tiež z: [https://www.linuxsecrets.com/elix-wiki/images/3/39/SecureOS\\_nakamura.pdf](https://www.linuxsecrets.com/elix-wiki/images/3/39/SecureOS_nakamura.pdf)
- [4] TINYPROXY. Dostupné tiež z: <https://github.com/tinyproxy/tinyproxy>
- [5] SHIROLIKAR, Abhijeet. Dostupné tiež z: <https://www.scionova.com/2019/04/08/securing-linux-with-mandatory-access-control/>
- [6] Keyboard shortcuts, 100 Keyboard Shortcuts to Use Linux Like a Pro, David Morelo. Dostupné tiež z: [https://linuxhint.com/100\\_keyboard\\_shortcuts\\_linux/](https://linuxhint.com/100_keyboard_shortcuts_linux/)
- [7] MAREŠ, Martin; BLACKHAM, Bernard. A New Contest Sandbox. *Olympiads in Informatics*. 2012, roč. 6. Dostupné tiež z: <https://www.privoxy.org/usermanual/actions-file.html>.
- [8] NAKAMURA, Yuichi. SELinux & AppArmor-Comparison of Secure OSes. Hitachi Software Engineering. 2007. Dostupné tiež z: [https://www.linuxsecrets.com/elix-wiki/images/3/39/SecureOS\\_nakamura.pdf](https://www.linuxsecrets.com/elix-wiki/images/3/39/SecureOS_nakamura.pdf)

# Kompresia multi-dimenzionálnych dát s využitím transformácie ich dimenzionality

<sup>1</sup>Heidar KHORSHIDIYEH, <sup>2</sup>Branislav MADOŠ, <sup>3</sup>Anton BALÁŽ

<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>heidar.khorshidiyeh@student.tuke.sk, <sup>2</sup>branislav.mados@tuke.sk,  
<sup>3</sup>anton.balaz@tuke.sk

**Abstrakt** — Článok je venovaný problematike reprezentácie geometrie viacrozmerných scén pomocou hierarchických dátových štruktúr založených na kvadrantových, resp. oktantových stromoch a z nich odvodených orientovaných acyklických grafoch. Cieľom tejto práce je overiť hypotézu, že je možné zvýšiť kompresiu údajov súvisiacich s geometriou viacrozmerných scén ich prevodom do scén s vyššou dimenzionalitou. S využitím algoritmu umožňujúceho transformovať geometriu scény na geometriu s vyššou, respektíve nižšou dimenzionalitou bolo na vybraných testovacích scénach overené, či je možné zvýšiť kompresný pomer zvýšením dimenzionality scény. Získané výsledky sú zhrnuté a vyhodnotené v záverečnej časti článku.

**Keywords** — morton order, space-filling curves, sparse voxel octrees, SVO, pointerless sparse voxel octrees, PSVO

## I. ÚVOD

Hierarchické dátové štruktúry (HDS) vo forme stromov a z nich odvodených orientovaných acyklických grafov – Directed Acyclic Graphs (DAG) je možné použiť na reprezentáciu viacrozmerných dát, napríklad dvojrozmerných obrazov v kartografii, technických výkresov v stavebnom priemysle, strojárstve, elektrotechnike atď., a tiež na reprezentáciu trojrozmerných scén vo filmoch, počítačových hrách a rôznych formách vizualizácií v oblasti vedy, výskumu a vývoja, medicíny a v mnohých ďalších oblastiach.

Jedným z typických prípadov ich použitia je využívanie oktantových stromov, alebo skrátene v angl. octrees a z nich odvodených orientovaných acyklických grafov na reprezentáciu geometrie trojrozmerných scén. V takom prípade pre každý voxel (to znamená objemový pixel – trojrozmerná analógia pixelu z dvojrozmerných obrázkov) v pravidelnej trojrozmernej mriežke voxelov, ktoré reprezentujú scénu, existuje jednobitová informácia, či je voxel aktívny alebo pasívny. Keď je voxel pasívny, je prázdny a je reprezentovaný bitom 0, keď je aktívny, je zaplnený a je reprezentovaný bitom 1. Aj keď je množstvo informácie o geometrii jedného voxelu veľmi malé, a je zaznamenané do iba jediného bitu, veľkosť trojrozmernej scény môže byť významná, rovnako ako jej prislúchajúca geometria. Napríklad, keď má scéna rozlíšenie  $1K^3$  ( $1024 \times 1024 \times 1024$ ) voxelov, celkové množstvo voxelov je  $2^{30}$ , rovnako ako počet bitov reprezentujúcich jej geometriu. Spotrebovávajú tak 1 Gb resp. 128 MB operačnej pamäte. Zdá sa to ako prijateľné množstvo, ale v prípade scény s rozlíšením  $64K^3$  ( $65536 \times 65536 \times 65536$ ) voxelov to znamená  $2^{48}$  voxelov a geometria scény v nekomprimovanom tvare v tom prípade zaberá 32 TB pamäte. Je preto nepraktické ukladať takéto objemy dát v nekomprimovanom tvare na sekundárne úložisko, a je prakticky nemožné ich reprezentovať v operačnej pamäti počítača alebo pamäti grafickej karty. Riešením je reprezentácia tohto druhu dát v komprimovanej podobe.

Pomocou bezstratovej kompresie dát s využitím špecifických komprimačných algoritmov a hierarchických dátových štruktúr je možné dosiahnuť kompresné pomery v rádoch  $10\,000 : 1$  a umožniť tak praktickú reprezentáciu aj rozsiahlych geometrií scén, ako je  $64K^3$  v operačnej pamäti počítača alebo v pamäti grafickej karty a umožniť vizualizáciu týchto scén.

Hlavným cieľom tejto práce je preskúmať, či je možné zvýšiť kompresný pomer dát, keď sú multi-dimenzionálne dáta s  $n$ -dimenziami reorganizované do dát s vyššou dimenziou. Testovaný bol preto rozdiel vo veľkosti reprezentácie tej istej obrazovej informácie v 2D a v 3D pri použití stromov neimplementujúcich vo svojej binárnej reprezentácii ukazovatele na dcérske uzly.

## II. SÚVISIACE PRÁCE

Hierarchické dátové štruktúry (HDS), ktoré sa používajú na uloženie geometrie dvojrozmerných obrazov používaných v kartografii alebo ako technické výkresy v elektrotechnike, strojárstve atď., a tiež na reprezentáciu trojrozmerných scén, sa vyvíjajú desaťročia. Boli vyvinuté Sparse Voxel Octrees, ktoré sú oktantovými stromami, pričom neumožňujú zlučovanie spoločných podstromov. Táto dátová štruktúra zahŕňa vo svojej implementácii ukazovatele na dcérske uzly, ale existuje aj verzia bez ich implementácie, ktorá sa nazýva Pointerless Sparse Voxel Octrees (PSVO) [1].

V roku 2013 bola vyvinutá dátová štruktúra s názvom High Resolution Sparse Voxel Directed Acyclic Graphs (SVDAG), ktorá umožňuje spájanie spoločných podgrafov. Táto dátová štruktúra bola zavedená v [2]. Každá časť vnútorného a listového uzla je v nej zarovnaná na 32 bitov, rovnako ako aj celková dátová štruktúra.

V roku 2016 bola v [3] zavedená dátová štruktúra SSVDAG (Sparse Voxel Directed Acyclic Graphs) zohľadňujúca symetriu. Umožňuje zlučovanie spoločných podgrafov nielen vtedy keď sú absolútne identické, ale tiež vtedy, keď je pre získanie ich identickosti potrebná transformácia zrkadlením v jednej alebo viacerých osiach súčasne. Listové uzly nie sú reprezentované ako klasický listový uzol dátovej štruktúry, ale ako mriežka voxelov v organizácii  $4 \times 4 \times 4$  voxelov.

V roku 2020 bola v [4] predstavená dátová štruktúra s názvom Pointerless Sparse Voxel Directed Acyclic Graphs. Kombinuje najlepšie vlastnosti dátovej štruktúry PSVO, keď neobsahuje ukazovatele na dcérske uzly, ktoré môžu predstavovať až viac ako 80% binárnej reprezentácie uzla. Na druhej strane umožňuje spájanie spoločných podstromov, ako dátová štruktúra SVDAG, a to aj bez ukazovateľov na dcérske uzly, vďaka implementácii návěstí (labels) a ich volaní (callers). Umožňuje tak ďalšie zvýšenie kompresného pomeru pri reprezentácii dát.

Témou súvisiacou s hierarchickými dátovými štruktúrami sú krivky vyplňajúce priestor [5]. V tejto práci sa používa Morton order krivka, ktorá sa tiež označuje ako Mortonov kód, Mortonova sekvencia alebo Z-order [6].

## III. NAVRHNUTÝ ALGORITMUS

Na uskutočnenie transformácie dimenzionality a vytvorenie hierarchickej dátovej štruktúry sú potrebné dva kroky:

- **Krok 1** Transformácia dimenzionality - transformácia n-rozmernej mriežky prvkov do m-rozmernej mriežky prvkov.
- **Krok 2** Konštrukcia stromov Sparse Voxel Trees (SVT) a Pointerless Sparse Voxel Trees (PSVT) ako hierarchickej dátovej štruktúry, ktorá predstavuje mriežku prvkov z kroku 1 v komprimovanej podobe.

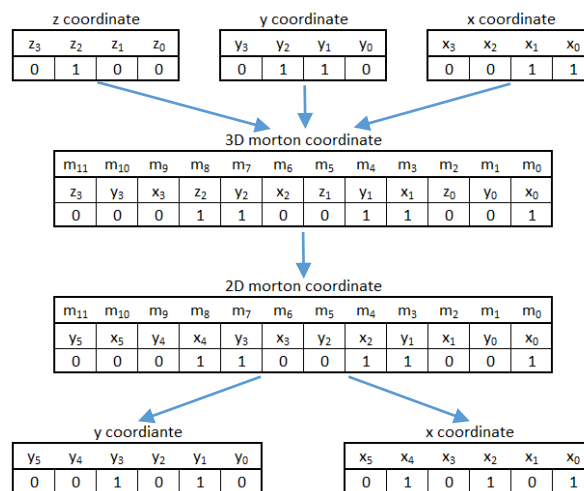
### A. Transformácia dimenzionality

Geometria n-rozmernej mriežky voxelov je štruktúra, kde každý voxel má svoje súradnice a je reprezentovaný 1 bitom. Keď je bit nastavený na 0, je voxel pasívny a keď je voxel nastavený na hodnotu 1 je aktívny. Dimenzionalita z 3D do 2D sa potom transformuje pomocou tohto trojkrokového algoritmu.

- **Krok 1** Ako je znázornené na Obrázku 1, súradnice x, y, z aktívnych voxelov 3D scény sa transformujú do jednej 3D Morton súradnice. Táto súradnica sa skonštruuje z trikrát viac bitov ako má každá zo súradníc x, y, z.
- **Krok 2** Hodnota 3D Morton súradnice predstavuje poradové číslo voxelu v linearizovanej forme pre voxel, ktorý má súradnice x, y, z v 3D mriežke voxelov. V tomto kroku sa binárna reprezentácia tejto súradnice deklaruje ako 2D Morton súradnica (s rovnakou binárnou reprezentáciou).
- **Krok 3** 2D Morton súradnica sa transformuje na súradnice x a y novej 2D mriežky pixelov.

Rovnaký postup je možné použiť aj na opačný smer transformácie napr. z 3D do 4D.

- **Krok 1** súradnice x, y, z aktívnych voxelov 3D scény sa transformujú do jednej 3D Morton súradnice.
- **Krok 2** hodnoty 3D Morton súradníc sa použijú ako hodnoty 4D Morton súradníc (s rovnakou binárnou reprezentáciou).
- **Krok 3** Vytvoria sa nové súradnice x, y, z, q voxelov v 4D mriežke zo 4D Morton súradnice.



Obr. 1 Súradnice obrazového prvku x, y, z sa transformujú do 3D Morton súradnice a potom sa interpretujú ako 2D Morton súradnica a nakoniec sa transformujú do súradníc obrazového prvku x, y.

### B. Konštrukcia hierarchickej dátovej štruktúry

Druhým krokom algoritmu je vytvorenie n-dimenzionálnej verzie hierarchickej dátovej štruktúry Sparse Voxel Octrees (označená ako SVT) v prípade, že sú začlenené ukazovatele na dcérske uzly, alebo n-dimenzionálnej verzie hierarchickej dátovej štruktúry Pointerless Sparse Voxel Octrees (označená ako PSVT) v prípade, že nie sú v dátovej štruktúre zastúpené žiadne ukazovatele na dcérske uzly. Ďalej sú popísané oba typy hierarchických dátových štruktúr, formalizované pomocou Backus-Naurovej Formy:

SVT ::= (n) <NODE>	PSVT ::= <NODE>
NODE ::= <INODE>   <LNODE>	NODE ::= <INODE>   <LNODE>
INODE ::= <CHNM> (p) <BIT> <PTS>	INODE ::= (2 <sup>D</sup> ) <HT>
LNODE ::= <CHNM> (q) <BIT>	LNODE ::= (2 <sup>D</sup> ) <BIT>
CHNM ::= (2 <sup>D</sup> ) <BIT>	HT = "0"   "1" <NODE>
PTS ::= (1)*(2 <sup>D</sup> ) <PT>	BIT ::= "0"   "1"
PT ::= (r) <BIT>	
BIT ::= "0"   "1"	

kde

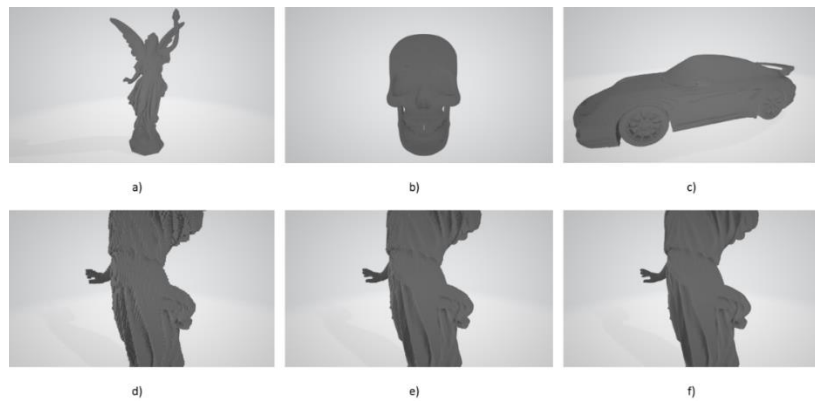
- p, q znamená počet rezervovaných bitov pre zarovnanie dĺžky
- r znamená počet bitov v ukazovateli na dcérsky uzol
- D znamená dimenzionalitu hierarchickej dátovej štruktúry

## IV. VÝSLEDKY TESTOV

Testy sa uskutočňovali pomocou voxelizovaných datasetov získaných procesom voxelizácie z troch polygonálnych povrchových modelov pôvodne uložených vo formáte súboru Wavefront OBJ. Týmito modelmi sú Angel Lucy (489 × 10<sup>3</sup> trojuholníkov), Skull (80 × 10<sup>3</sup> trojuholníkov) a Porsche (22 × 10<sup>3</sup> trojuholníkov). Tieto modely pozostávajú z trojuholníkov, ktoré boli v procese voxelizácie transformované na aktívne voxely. Rozlíšenie, do ktorých boli modely voxelizované, je 128<sup>3</sup>, 256<sup>3</sup>, 512<sup>3</sup>, 1K<sup>3</sup> (1024<sup>3</sup>), 2K<sup>3</sup> (2048<sup>3</sup>), 4K<sup>3</sup> (4096<sup>3</sup>). Celkový počet modelov, keď sa zohľadní aj rozlíšenie scény, je teda 18. Ich vizualizácie ukazujú Obr. 2 a parametre zhŕňa Tab. 1.

Podiel aktívnych voxelov v týchto scénach sa pohybuje od 28,95% pre model Skull, ktorý bol voxelizovaný na rozlíšenie 128<sup>3</sup>, až po 0,26% pre model Angel Lucy, ktorý bol voxelizovaný na rozlíšenie 4096<sup>3</sup> (4K<sup>3</sup>). Absolútny počet aktívnych voxelov je naopak najmenší v prípade najnižšieho rozlíšenia, keď model Angel Lucy, ktorý bol voxelizovaný na rozlíšenie 128<sup>3</sup>, pozostáva z 22,48 tisíc aktívnych voxelov a najväčší je v prípade modelu Skull, ktorý bol voxelizovaný na rozlíšenie 4096<sup>3</sup> (4K<sup>3</sup>), kde je to až 64,61 mil. aktívnych voxelov.





Obr. 2. Vizualizácia voxelizovaných scén použité pre testovacie účely: a) Angel Lucy 512<sup>3</sup>, b) Skull 512<sup>3</sup>, c) Porsche 512<sup>3</sup>, d) detail modelu Angel Lucy 256<sup>3</sup>, e) detail modelu Angel Lucy 512<sup>3</sup>, f) detail modelu Angel Lucy 1K<sup>3</sup>.

Tab. 1 Charakteristiky volumetrických datasetov, ktoré boli vytvorené v procese voxelizácie polygonálnych modelov uložených vo WaveFront OBJ súborovom formáte. Tabuľka sumarizuje počet aktívnych voxelov a ich percentuálny podiel na celkovom počte voxelov v datasete, pre každý model a rozlíšenie.

	Rozlíšenie					
Rozlíšenie	128 <sup>3</sup>	256 <sup>3</sup>	512 <sup>3</sup>	1024 <sup>3</sup>	2048 <sup>3</sup>	4096 <sup>3</sup>
Voxely [mil]	2	16	128	1 024	8 192	65 536
<b>Angel Lucy - 488 880 tri</b>						
Akt. Voxely [tis]	22,48	91,52	366,58	1 453,10	5 685,86	21 656,43
Percento	8,78	4,47	2,24	1,11	0,54	0,26
<b>Skull - 80 016 tri</b>						
Akt. Voxely [tis]	74,10	298,85	1 192,04	4 688,08	17 958,71	64 608,51
Percento	28,95	14,59	7,28	3,58	1,71	0,77
<b>Porsche - 22011 tri</b>						
Akt. Voxely [tis]	54,20	233,04	969,11	3 938,35	15 539,54	58 673,98
Percento	21,17	11,38	5,91	3,00	1,48	0,70

Výsledky vytvorenia Pointerless Sparse Voxel Quadrees (PSVQ) hierarchickej dátovej štruktúry sú uvedené v Tab. 2. Počet uzlov sa zvyšuje predpokladaným spôsobom a rovnako tak aj veľkosť súboru v MB, keď je každý uzol reprezentovaný iba 4 bitmi Child Node Mask (nibble), pretože v tomto HDS sú iba štyri možné dcérske uzly. Nie sú implementované ukazovatele na dcérske uzly a nepoužívajú sa žiadne rezervované bity.

Tab. 2. Výsledky zostavenia hierarchickej dátovej štruktúry založené na Pointerless Sparse Voxel Quadrees (PSVQ) na rôznych modeloch a veľkostiach ich voxelizácií, v počte uzlov a vo veľkosti súborov v megabajtoch (MB).

PSVQ	Rozlíšenie					
	128 <sup>3</sup>	256 <sup>3</sup>	512 <sup>3</sup>	1024 <sup>3</sup>	2048 <sup>3</sup>	4096 <sup>3</sup>
<b>Angel Lucy</b>						
[uzly]	14229	60515	248754	1001341	3979476	15592897
[MB]	0,01	0,03	0,12	0,48	1,90	7,44
<b>Skull</b>						
[uzly]	50173	205836	832834	3328156	13109355	50295715
[MB]	0,02	0,10	0,40	1,59	6,25	23,98
<b>Porsche</b>						
[uzly]	32630	148438	642405	2687781	10974405	43493740
[MB]	0,02	0,07	0,31	1,28	5,23	20,74

Boli tiež vytvorené verzie Sparse Voxel Quadtrees (SVQ) s implementovanými ukazovateľmi, pričom 4 bity masky dcérskeho uzla sú doplnené 28 ďalšími bitmi, aby bola táto časť uzla zarovnaná na 32 bitov. Každý ukazovateľ na dcérsky uzol je tiež dlhý 32 bitov. Výsledky v počte uzlov a celkovej veľkosti konkrétneho SVQ v MB sú zhrnuté v Tab. 3.

Tab. 3. Výsledky zostavenia hierarchickej dátovej štruktúry založenej na Sparse Voxel Quadtrees (SVQ) na rôznych modeloch a veľkostiach ich voxelizácií, v počte uzlov a vo veľkosti súborov v megabajtoch (MB).

SVQ	Rozlíšenie					
	128 <sup>3</sup>	256 <sup>3</sup>	512 <sup>3</sup>	1024 <sup>3</sup>	2048 <sup>3</sup>	4096 <sup>3</sup>
<b>Angel Lucy</b>						
[uzly]	14229	60515	248754	1001341	3979476	15592897
[MB]	0,11	0,46	1,90	7,64	30,36	118,96
<b>Skull</b>						
[uzly]	50173	205836	832834	3328156	13109355	50295715
[MB]	0,38	1,57	6,35	25,39	100,02	383,72
<b>Porsche</b>						
[uzly]	32630	148438	642405	2687781	10974405	43493740
[MB]	0,25	1,13	4,90	20,51	83,73	331,83

Výsledky vytvorenia Pointerless Sparse Voxel Octrees (PSVO) hierarchickej dátovej štruktúry sú uvedené v Tab. 4. Počet uzlov sa zvyšuje predpokladaným spôsobom a rovnako tak aj veľkosť súboru v MB, keď je každý uzol reprezentovaný iba 8-bitovou maskou Child Node Mask, pretože v tomto HDS je osem možných dcérskych uzlov. Na dcérske uzly nie sú vytvárané žiadne ukazovatele a nepoužívajú sa žiadne vyhradené bity.

Verzie Sparse Voxel Octrees (SVO) s implementovanými ukazovateľmi boli tiež vytvorené, pričom bolo použitých 8 bitov masky uzla a 24 ďalších bitov bolo doplnených, aby bola táto časť uzla zarovnaná na 32 bitov. Každý ukazovateľ na dcérsky uzol je tiež dlhý 32 bitov. Výsledky v počte uzlov a celkovej veľkosti konkrétneho SVO v MB sú zhrnuté v Tab. 5.

Ako je zrejmé z uvedených tabuliek, veľkosť PSVO je oproti PSVQ menšia, od 2% do 8%, avšak táto medzera sa znižuje s nárastom rozlíšenia voxelizácie. Veľkosť SVO je podstatne menšia oproti SVQ, od 51% do 54%, a táto medzera sa s nárastom rozlíšenia voxelizácie veľmi mierne znižuje.

Tab. 4. Výsledky zostavenia hierarchickej dátovej štruktúry založenej na Pointerless Sparse Voxel Octrees (PSVO) na rôznych modeloch a veľkostiach ich voxelizácií v počte uzlov a veľkosti súborov uvedená v megabajtoch (MB).

PSVO	Rozlíšenie					
	128 <sup>3</sup>	256 <sup>3</sup>	512 <sup>3</sup>	1024 <sup>3</sup>	2048 <sup>3</sup>	4096 <sup>3</sup>
<b>Angel Lucy</b>						
[uzly]	6847	29327	120843	487423	1940523	7626380
[MB]	0,01	0,03	0,12	0,46	1,85	7,27
<b>Skull</b>						
[uzly]	23822	97924	396768	1588811	6276893	24235606
[MB]	0,02	0,09	0,38	1,52	5,99	23,11
<b>Porsche</b>						
[uzly]	14964	69166	302205	1271317	5209667	20749206
[MB]	0,01	0,07	0,29	1,21	4,97	19,79

Tab. 5 Výsledky zostavenia hierarchickej dátovej štruktúry založenej na Sparse Voxel Octrees (SVO) na rôznych modeloch a veľkostiach ich voxelizácií v počte uzlov a veľkosť súborov uvedená v megabajtoch (MB).

SVO	Rozlíšenie					
	128 <sup>3</sup>	256 <sup>3</sup>	512 <sup>3</sup>	1024 <sup>3</sup>	2048 <sup>3</sup>	4096 <sup>3</sup>
<b>Angel Lucy</b>						
[uzly]	6847	29327	120843	487423	1940523	7626380
[MB]	0,05	0,22	0,92	3,72	14,80	58,18
<b>Skull</b>						
[uzly]	23822	97924	396768	1588811	6276893	24235606
[MB]	0,18	0,75	3,03	12,12	47,89	184,90
<b>Porsche</b>						
[uzly]	14964	69166	302205	1271317	5209667	20749206
[MB]	0,11	0,53	2,31	9,70	39,75	158,30

## V. ZÁVER

Článok sa zaoberal problematikou použitia hierarchických dátových štruktúr pre účely reprezentácie geometrie viacrozmerých scén s cieľom preskúmať hypotézu, že zvýšenie dimenzionality pravidelnej mriežky obrazových prvkov môže priniesť vyšší kompresný pomer pri komprimovanej reprezentácii geometrie takejto scény. Zostavený bol algoritmus pre zmenu dimenzionality scény tak, aby umožňoval jej transformáciu z 2D na 3D prípadne 4D a naopak a následne bol aplikovaný algoritmus pre kódovanie viacrozmernej scény do stromovej hierarchickej dátovej štruktúry. Bola použitá hierarchická dátová štruktúra podobná PSVO resp. SVO s modifikáciami pre 2D a 3D.

Testy sa uskutočňovali na troch rôznych polygonálnych povrchových modeloch Angel Lucy, Skull a Porsche, ktoré boli zo svojho pôvodného zobrazenia v modeli WaveFront OBJ voxelizované do šiestich rôznych rozlíšení od 128<sup>3</sup> do 4096<sup>3</sup> voxelov. Výsledky testov ukazujú, že u všetkých modelov a ich rozlíšení bol model reprezentovaný v trojrozmernej reprezentácii úspornejší, z hľadiska priestoru ktorý jeho binárna reprezentácia zaberá, ako v prípade dvojrozmernej reprezentácie tohoto modelu. Vo verzii bez implementácie ukazovateľov na dcérske uzly bol rozdiel vo veľkosti ich binárnej reprezentácie relatívne malý, vo veľkosti 2 až 8%. Rozdiel sa s rastom rozlíšenia zmenšoval. Vo verzii s implementovanými ukazovateľmi na dcérske uzly bol rozdiel oveľa výraznejší, na úrovni medzi 51% až 54% v prospech 3D verzie reprezentácie, pričom tieto výsledky boli dosiahnuté pre všetky modely a všetky ich rozlíšenia.

## POĎAKOVANIE

Tento výskum bol podporený projektom 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market.

## REFERENCIE

- [1] B. Madoš, N. Ádám a M. Štancel, "Representation of Dense Volume Datasets Using Pointerless Sparse Voxel Octrees With Variable and Fixed-Length Encoding", 2021. In: 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI): 21. - 23.1.2021: Herľany, Slovakia, s. 343-348, Košice, Slovakia: Institute of Electrical and Electronics Engineers, 2021 ISBN 978-1-7281-8054-0.
- [2] V. Kämpe, E. Sintorn a U. Assarsson, "High Resolution Sparse Voxel DAGs". In: ACM Trans. Graph. 32.4 (júl 2013). ISSN: 0730-0301.doi:10.1145/2461912.2462024.71.
- [3] A. J. Villanueva, F. Marton a E. Gobbetti, "SSVDAGs Symmetry-Aware Sparse Voxel DAGs". In: Proceedings of the 20th ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games. I3D '16. Redmond, Washington: Association for Computing Machinery, 2016, s. 7-14. ISBN: 9781450340434. doi:10.1145/2856400.285642.
- [4] L. Vokorokos, B. Madoš a Z. Bilanová, "PSVDAG: Compact Voxelized Representation of 3D Scenes Using Pointerless Sparse Voxel Directed Acyclic Graphs". In: Computing and Informatics (jan. 2020), s. 587-616. doi:10.31577/cai\_2020\_3\_587.
- [5] A. Laszloffy, J. Long, a A. K. Patra, "Simple data management, scheduling and solution strategies for managing the irregularities in parallel adaptive finite element simulations", Parallel Computing, 26, 1765-1788.
- [6] G.M. Morton, "A Computer Oriented Geodetic Data Base and a New Technique in File Sequencing", Research Report. International Business Machines Corporation (IBM), Ottawa, Canada, 20.

# Návrh nekonvenčného zdroja energie na konvenčné účely

<sup>1</sup>Miroslav ANDRAŠ, <sup>2</sup>Dušan MEDVEĎ

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

<sup>1</sup>miroslav.andras@student.tuke.sk, <sup>2</sup>dusan.medved@tuke.sk

**Abstrakt** — Tento príspevok sa zaoberá návrhom a skonštruovaním jednoduchého nekonvenčného modelu ohrevu vody, ktorý môže prispieť k zníženiu finančných výdavkov v domácnosti. Model ohrevu vody, ktorý je súčasťou tohto článku, využíva obnoviteľný zdroj energie, akým je slnečné žiarenie, ktoré sa následne transformuje na elektrickú energiu pomocou fotovoltaických panelov, ktoré napájajú ohrevnú špirálu, slúžiacu pre ohrev vody. Ako ďalší zdroj energie pre ohrev je použitý priamy ohrev vody, a to za pomoci polyetylénového vodovodného potrubia, ktoré pracuje na princípe konvenčných kolektorových panelov. Keďže v skonštruovanom modeli boli použité použité fotovoltaické panely, tak článok obsahuje aj diagnostiku použitých fotovoltaických panelov (termokamerou), a tým ozrejmuje ich aktuálny stav. Príspevok ďalej popisuje jednotlivé súčasti, ktoré boli potrebné pre skonštruovanie modelu a aj výsledné parametre, ako sú energetické bilancie jednotlivých zdrojov energie pre ohrev, výsledky merania a možnú finančnú úsporu pri kombinovanej prevádzke navrhnutého modelu a konvenčného plynového ohrevu vody.

**Kľúčové slová** — fotovoltaický systém, kolektorový systém, ohrev vody, finančná úspora

## I. ÚVOD

Energetická kríza sa postupom času vo svete začína čoraz viac prejavovať. V priebehu 21. storočia budú spotrebované väčšinové zásoby fosílnych palív, ktoré prevládajú vo svete ako zatiaľ lacný zdroj energie. Tento druh palív je využívaný aj ako zdroj energie v domácnostiach, či už pre výrobu tepla alebo elektrickej energie. V súčasnej dobe, kedy environmentálna problematika priamo súvisí s problematikou úspor energie v domácnostiach, sa okrem samotnej úspory energie kladie dôraz aj na environmentálnu stránku, a to spôsobom použitia šetrnejších zdrojov energie, ako napríklad solárne systémy používané priamo v domácnostiach. Preto sa aj tento článok bude zameriavať na úsporu energie pomocou solárnych systémov, slúžiacich na ohrev vody v domácnosti. V tomto článku sa budeme hlavne zameriavať na úsporu energie v domácnostiach použitím fotovoltaického (FV) a kolektorového systému. Súčasťou článku bude aj praktický návrh samotného riešenia úspor energie v domácnostiach, a to za pomoci FV systému, ktorého súčasťou budú použité FV panely, ktoré boli používané vo FV elektrárni. Tento FV systém bude použitý na ohrev vody, súčasťou ohrevu bude aj jednoduchý ohrev fungujúci na princípe solárnych kolektorov. Článok sa zaoberá, o.i. súčasnou problematikou obnoviteľných zdrojov energie, teoretickými poznatkami z oblasti FV a kolektorových systémov, ich výhody, nevýhody a typy. Ďalej je článok zameraný na praktický návrh a skonštruovanie modelu pre ohrev pitnej alebo úžitkovej vody v domácnosti a v závere sú prezentované namerané výsledky navrhnutého modelu ohrevu vody a celková finančná úspora.

## II. SÚČASNÝ STAV PROBLEMATIKY

V súčasnej dobe sa viac kladie dôraz na úsporu energie či v domácnostiach alebo v priemysle, a to hlavne z dôvodu zvyšujúcich sa cien fosílnych palív, ktoré prevládajú ako energetické zdroje v celosvetovej energetike. Fosílna paliva sú napr. ropa, zemný plyn, čierne alebo hnedé uhlie, pričom v celkovej spotrebe prevláda uhlie (80%), nasleduje ropa (10%) a zemný plyn (10%). Hlavným problémom fosílnych palív je, že sú vyčerpatelné a v súčasnej dobe sa predpokladá, že v priebehu 21. storočia sa vyčerpajú väčšinové zásoby týchto palív; uhlie v priebehu 312 rokov, ropa v priebehu 41 rokov a zemný plyn v priebehu 66 rokov. Fosílna palivá majú veľký dopad na životné prostredie a to hlavne spaľovaním palív, pri ktorom sa vytvárajú sírové, uhličité a dusíkové kyseliny, ktoré sa následne na zemský povrch dostávajú v podobe kyslých dažďov, ktoré nepriaznivo pôsobia na životné prostredie.

Východiskom z tejto situácie sú obnoviteľné zdroje energie, ktoré sa zatiaľ vo väčšine považujú za netradičné zdroje energie. Pri čerpaní týchto zdrojov sa využívajú nové technológie, ktoré sú zatiaľ v celosvetovej produkcii energie malo používané a stále sa pracuje na ich zdokonaľovaní. Obnoviteľné zdroje predstavujú ekologicky čistý zdroj energie, ktorý by dokázal pokryť celosvetovú spotrebu energie. No hlavne z ekonomického hľadiska, nie je možné, aby sa výroba sústredila len na obnoviteľné zdroje energie. Obnoviteľné zdroje sú, na rozdiel od fosílnych palív, šetrné k životnému prostrediu a aj preto vznikajú rôzne programy, ktoré podporujú väčšie využitie obnoviteľných zdrojov v domácnostiach. Najčastejšími využívanými obnoviteľnými zdrojmi v domácnostiach sú fotovoltické panely, solárne kolektory alebo tepelné čerpadlá. Nevýhodou sú veľké počiatočné finančné náklady, no vďaka práve týmto systémom je možné dôjsť k výraznej úspore energie v domácnostiach.

### III. NÁVRH A SKONŠTRUOVANIE NEKONVENČNÉHO SOLÁRNEHO OHREVVU VODY V DOMÁCNOSTI

V súčasnej dobe je ako hlavný zdroj ohrevu pitnej alebo úžitkovej vody v domácnostiach použitý ohrev, ktorý využíva ako zdroj energie fosílnu palivú, čo má nepriaznivý vplyv na životné prostredie a aj na ekonomickú stránku, keďže ceny fosílnych palív každým rokom stúpajú. Z tohto dôvodu sme sa rozhodli pre návrh a skonštruovanie jednoduchého, lacného a účinného energetického zdroja pre ohrev vody v domácnosti. Ako primárny zdroj energie bude využívať solárnu energiu, ktorá je finančne výhodná a patri medzi obnoviteľné zdroje energie. Tento model kombinuje základné systémy, ktoré sa využívajú vo svete pre transformáciu solárnej energie na iný druh energie. Prvým z týchto systémov je fotovoltika, ktorá využíva fotoelektrický jav pre transformáciu solárnej energie na elektrickú. Druhým systémom je kolektorový systém, ktorý využíva priame slnečné žiarenie, ktoré dopadá na absorbér a odovzdáva energiu vo forme tepla a následne toto teplo slúži pre ohrev kvapaliny cirkulujúcej vo vnútri absorbéra. V navrhnutom modeli ohrevu vody sú použité FV panely, ktoré priamo napájajú elektrickou energiou elektrickú špirálu, o výkone 1600 W. Táto špirála je prispôbená pre prevádzku v konvenčných bojleroch, preto je ideálna pre použitie v predmetnom modeli. Celkový predpokladaný výkon FV panelov je pri ideálnych podmienkach 630 Wp, pričom tento výkon je možné dosiahnuť len pri nových FV paneloch. V tomto modeli boli použité (čiastočne poškodené) FV panely, ktorých parametre v závislosti od zapojenia, sú nasledovné: napätie naprázdno 175 V a elektrický prúd dosahuje hodnotu maximálne 2,5 A, pričom ide o jednosmernú elektrickú sústavu. Tieto hodnoty boli maximálne hodnoty, ktoré boli namerané počas merania 10.5.2021 čo znamená, že výkon FV panelov môže byť aj vyšší v nasledujúcich mesiacoch, keďže výkon kolíše v závislosti od vonkajších vplyvov, ako je napríklad teplota, vlhkosť vzduchu alebo oblačnosť, ktorá priamo súvisí s dopadajúcim slnečným žiarením. Elektrická špirála je uložená vo výmenníku tepla, ktorého kapacita je 100 litrov. Solárny systém, ako prídavný zdroj energie pre ohrev vody, využíva priame slnečné žiarenie, v podobe cirkulácie vody v LDPE rúrach, ktoré v praxi slúžia pre prepravu pitnej alebo úžitkovej vody. V tomto solárnom systéme slúžia ako jednoduchý absorbér slnečného žiarenia. Dopadajúce slnečné žiarenie na potrubie odovzdá energiu v podobe tepla, ktoré následne využijeme na ohrev pitnej a úžitkovej vody cirkulujúcej v danom systéme. Ide o jednoduchý princíp, ktorý využívajú napr. aj certifikované továrenské ploché solárne kolektory, samozrejme s vyššou účinnosťou.

#### A. Jednoduchý kolektorový ohrev v navrhnutom modeli

Pre ohrev vody v navrhnutom nekonvenčnom modeli, je použité vodárenské potrubie vyrobené z polyetylénu o priemere 32 mm. V tomto systéme zohráva jednoduchú úlohu, a to ako ohrev vody, kedy voda privádzaná do zásobníka, v ktorom je uložené výhrevné teleso napájané z FV panelov, sa najprv predhreje vo vodovodnom polyetylénovom potrubí a následne, po zopnutí kontaktov termostatu, je predhriata voda precirkulovaná za pomoci obehového čerpadla privádzaná do 100-litrového zásobníka vody, kde na zvyšný ohrev vody slúži elektrická ohrevná špirála. Princíp takéhoto ohrevu spočíva v konverzii slnečného žiarenia, dopadajúceho na polyetylénové potrubie, ktoré slúži ako absorbér slnečnej energie, ktorá je potom využitá na ohrev vody cirkulujúcej vo vnútri potrubia. Na podobnom princípe fungujú aj klasické továrenské solárne kolektory. Výhodou takéhoto ohrevu je jednoduchosť a nízke náklady na výrobu, nevýhodou je nízka účinnosť, ktorá je približne 35 až 40 %.



Obr. 1 Jednoduchý model kolektorového ohrevu vody

#### *B. Potrebné komponenty, pre návrh modelu ohrevu vody*

Ako jeden zo zdrojov energie pri navrhnutom modeli ohrevu vody v domácnosti, boli použité FV panely značky **Q.SMART UF L 105** od nemeckého výrobcu Q.cells, ktorý je jedným z popredných výrobcov FV panelov v Európe. FV panel Q.SMART UF L 105 je polykryštalický, bezrámový, tenkostenný FV panel, ktorý je vyrobený z tvrdeného skla. Počet FV článkov v jednom paneli je 100 ks. FV panely Q.SMART UF L 105 využívajú technológiu CIGS (Copper-Indium-Gálium-selenidu), to zvyšuje účinnosť získania energie z viacerých spektier slnečného žiarenia, pričom využíva aj žiarenie mimo viditeľného spektra a premieňa ho na elektrinu. Panely využívajúce technológiu CIGS majú v zamračenom a daždivom počasí až o 20 až 30% vyšší výkon v porovnaní s monokryštalickými FV panelmi s rovnakými nominálnymi hodnotami. V priebehu dňa tak pracujú dlhší čas, a to aj v zhoršených svetelných alebo náročnejších klimatických podmienkach, a sú preto vhodné k celoročnému využívaniu. Sú z oboch strán sklenené čo zvyšuje ich životnosť, ktorá je 30 až 40 rokov. Účinnosť panelov tohto typu je 11,17%. Tieto FV panely spĺňajú bezpečnostný štandard podľa EIC 61730 a spadajú do bezpečnostnej triedy II.



Obr. 2 Fotovoltický panel Q.SMART UF L 105

#### **Zásobník vody**

Na obrázku nižšie je možné vidieť nádrž na vodu, ktorá v navrhnutom modeli použitá ako zásobník teplej vody, pričom v zásobníku je nainštalovaná elektrická ohrevná špirála, ktorá slúži na ohrev pitnej vody v sústave, takže dá sa povedať, že ide o zjednodušený princíp ohrevu, ako pri klasických továrenských bojleroch. Zásobník má kapacitu 100 litrov a je vyrobený z plechových formátov valcovaných za tepla, pričom hrúbka plechu dosahuje 4 mm. Rozmery zásobníka sú 70 cm na dĺžku, 51 cm na šírku a 70 cm na výšku. V zásobníku sú inštalované

3 vývody, jeden pre prívod studenej vody do zásobníka a dva vývody pre odvod teplej vody, pričom jeden slúži aj pre inštaláciu termostatu a meranie teploty vody v sústave. Celý zásobník je tepelne izolovaný pomocou izolačnej vrstvy.



Obr. 3 Zásobník teplej vody

### Ohrevná špirála Tatramat 1600 W 5113/2

Pre ohrev vody je použitá ohrevná špirála, ktorá je nainštalovaná v spodnej časti zásobníka vody. Špirála tohto typu je využívaná v klasických konvenčných bojleroch, využívaných pre ohrev vody, takže je prispôbená pre prevádzku vo vode. Jediné časti, ktoré sú odkryté, sú kontakty pre napájanie špirály. Veľmi dôležité je, aby pri prevádzke bola špirála vždy pod vodou a aby nedošlo k jej poškodeniu spôsobenom prehriatím. Výhrevná špirála je vyrobená pre striedavé maximálne napájacie napätie 250 V a maximálny príkon špirály je 1600 W. V navrhnutom modeli ohrevu vody je špirála napájaná jednosmerným napätím, o hodnote maximálne 175 V, pri maximálnom dodávanom elektrickom výkone z FV panelov, pričom výkon výhrevnej špirály je závislý od intenzity dopadajúceho slnečného žiarenia na FV panely.

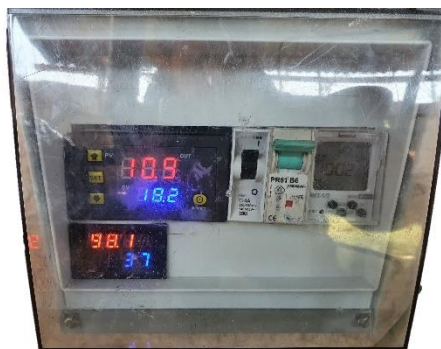


Obr. 4 Ohrevná elektrická špirála Tatramat 1600W

### Meranie a ovládanie pre FV systém

V navrhnutom modeli, ktorý ako zdroj energie využíva elektrickú energiu vyrobenú za pomoci FV panelov, je súčasťou modelu aj elektrický rozvádzač, ktorý slúži pre istenie, meranie a riadenie elektrických komponentov použitých v modeli. V elektrickom rozvádzači sú pre istenie použité dva ističe, prvý istič B6 od značky SEZ slúži pre istenie elektrického obvodu, ktorý napája elektrické zariadenia ako termostat, meracie zariadenie FV systému, obehové čerpadlo a digitálne spínacie hodiny. Ďalší istič G4A od značky SIEMENS, slúži pre istenie ohrevnej elektrickej špirály uloženej v zásobníku, pričom špirála je napájaná z FV panelov. Použitím istiaceho prvku v tomto obvode sa okrem istenia zamedzí pri spínaní a rozpínaní aj vzniku veľkého elektrického oblúka, spôsobeného jednosmerným napätím napájaného elektrickú špirálu, keďže istič obsahuje zhášaciu komoru. Rozvádzač obsahuje okrem istiacich prvkov aj termostat značky Ketotek, ktorý je napájaný napätím AC 230 V. Obsahuje jeden spínací kontakt s maximálnym možným prúdovým zaťažením 10 A. V navrhnutom modeli slúži pre spínanie obehového čerpadla, pokiaľ teplota vody v zásobníku presiahne 45 °C, spustí sa obehové čerpadlo, ktoré precirkuluje ohriatu vodu. Pre meranie elektrického výkonu dodávaného FV panelmi je použité meracie zariadenie od značky Ketotek, obsahujúce voltmeter a ampérmeter. Na voltmeter je možné pripojiť maximálne napätie nie vyššie ako DC 200 V a maximálny možný

elektrický prúd prechádzajúci obvodom nesmie byť viac ako 10 A. Meracie zariadenie je napájané napätím DC 12 V. V rozvádzači sú inštalované aj digitálne spínacie hodiny, slúžiace pre vypínanie systému merania a riadenia vo večerných a nočných hodinách, kedy nedochádza k ohrevu vody.



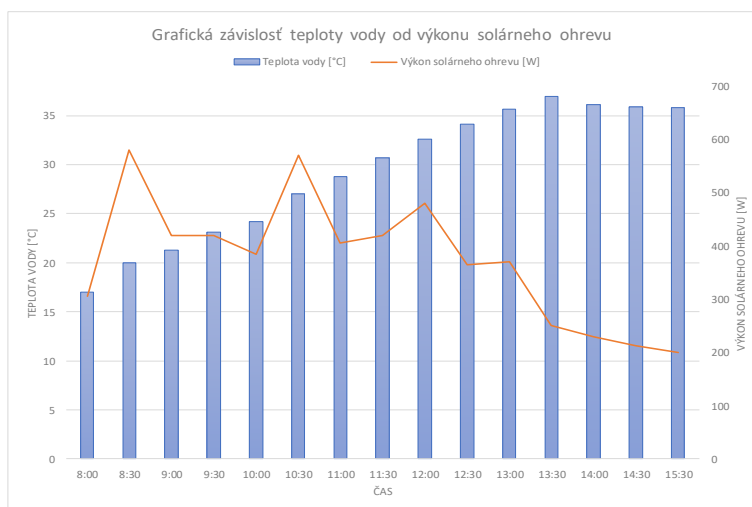
Obr. 5 Rozvádzač pre meranie a ovládanie FV systému

#### IV. MERANIE VÝKONNOSTI NAVRHNUTÉHO NEKONVENČNÉHO OHREVVU VODY

##### A. Meranie 11.5.2021

Tabuľka 1  
 Namerané hodnoty teploty vody a výkonu solárneho ohrevu dňa 11.2.2021

Čas	8:00	8:30	9:00	9:30	10:00	10:30	11:00	11:30	12:00	12:30	13:00	13:30	14:00	14:30	15:30
Teplota vody [°C]	17	20	21,3	23,1	24,2	27	28,8	30,7	32,6	34,1	35,7	37	36,1	35,9	35,8
Výkon solárneho ohrevu [W]	305	580	420	420	385	570	405	420	440	365	370	250	230	212	200



Obr. 6 Grafické znázornenie výkonu solárneho ohrevu a teploty vody v zásobníku 11.5.2021

Meranie prebiehalo 11.5.2021 od 8:00 do 15:00, priemerná pocitová teplota počas dňa bola 24°C, pričom počasie bolo polooblačné. Počiatočná teplota vody o 8:00 dosahovala teplotu 17°C a najvyššia dosiahnutá teplota vody v zásobníku o 13:30 dosahovala teplotu 37°C.

##### B. Výpočet finančnej úspory

Pokiaľ je v domácnosti použitý plynový ohrev vody, tak pri dnešných cenách plynu okolo 0,06 €/kWh to znamená, že ak v domácnosti žijú 2 osoby a spotreba teplej vody je 90 litrov, pričom plynový kotol ohrieva vodu z 10°C na 45°C a pritom spotrebuje 3,9 kWh, čo je denne 0,23 €, tak za pol roka to je 42 €. Pokiaľ by bol navrhnutý model rozoberaný v tomto článku použitý pre kombinovaný ohrev vody, spolu s plynovým ohrevom, celková denná finančná úspora by bola 0,18 €, celkovo za pol roka by bolo možné ušetriť 32 € v období od mesiaca apríl do mesiaca október. Samozrejme, výkon navrhnutého modelu ohrevu vody je závislý od intenzity dopadajúceho slnečného žiarenia. Pri výpočte tejto úspory bolo zahrnuté meranie z dňa 10.5.2021, kedy priemerný výkon navrhnutého modelu ohrevu vody bol počas dňa 434 W, čo je 3,08 kWh.



## V. ZÁVER

Tento článok bol zameraný na návrh nekonvenčného ohrevu vody vyžívaného v domácnosti, ktorý pre ohrev vody využíva solárnu energiu. Dôležitým faktorom je teda poloha navrhnutého modelu ohrevu vody, ktorý je situovaný na východnom Slovensku v oblasti Prešova, kedy počas merania v období mesiaca máj, bola energia dopadajúceho slnečného žiarenia  $4,74 \text{ kWh}\cdot\text{m}^{-2}$ . V modeli bol použitý jednoduchý ohrev vody, a to za pomoci polyetylénového potrubia, ktoré slúžilo ako absorbér a zachytávalo tepelnú energiu zo slnečného žiarenia, a tým ohrievalo vodu cirkulujúcu vo vnútri potrubia. Pri výpočte energetickej bilancie navrhnutého systému bola vypočítaná celková plocha polyetylénového potrubia  $3,02 \text{ m}^2$ . Na základe výpočtu bol celkový výkon slnečného žiarenia dopadajúceho na plochu polyetylénového potrubia stanovený na  $1,1 \text{ kW}$ , a to v období mesiaca máj. No účinnosť takéhoto jednoduchého modelu kolektora, na základe výpočtu z energetickej bilancie, bol stanovený na približne  $25 \%$ , čo je výkon okolo  $250 \text{ W}$  počas jasného počasia v mesiaci máj. Samozrejme, keďže kvôli absencii solarimetra nebolo možné presne stanoviť hodnotu dopadajúceho slnečného žiarenia, tak tieto výsledky sú aproximované na základe výpočtu z energetickej bilancie. Pri navrhnutom modeli boli použité aj FV panely, ktoré priamo napájajú ohrevnú špirálu uloženú v zásobníku vody, keďže panely nie sú nové a majú známky poškodenia, ako praskliny ochranného skla alebo delaminácia, bola súčasťou tohto článku aj optická diagnostika stavu FV panelov, z čoho vyplýva, že hoci sú FV panely poškodené, tak napriek tomu majú dostatočný výkon na to, aby ich bolo možné opätovne použiť. Konkrétne v tomto jednoduchom nekonvenčnom modeli ohrevu vody vyžívaného v domácnosti, ktorého používaním v kombinácii s klasickým konvenčným ohrevom, ako napríklad plynovým, je možné docieľiť až  $70 \%$  finančnej úspory pre domácnosť s dvomi osobami, v období od mesiaca apríl do mesiaca október. Preto tento model jednoduchého nekonvenčného ohrevu vody považujeme za vhodnú a lacnú alternatívu pre úsporu energie v jarnom až jesennom období. Nevýhodou tohto navrhnutého systému je, že v zimnom období nie je možnosť použitia tohto ohrevu vody, pretože by mohlo dôjsť k jeho poškodeniu spôsobenému zamrznutím vody v systéme, keďže tam nie je použitá teplovýmenná nemrznúca kvapalina, ktorá je využívaná v konvenčných systémoch pre ohrev vody.

## POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a ďalej Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu SR a SAV podporenej grantom VEGA 1/0757/21.

## LITERATÚRA

- [1] Kolcun, M., Džmura, J., Mešter, M., Pavlík, M.: *Elektrárne*. 2.vyd. Košice: TUKE, 2017 ISBN 978-80-971747-2-9.
- [2] *Slovenské fotovoltické elektrárne*: [online], [cit. 2021-15-05], dostupné na internete: <<https://www.seas.sk/fotovolticke-elektrarne>>
- [3] Andraš, M.: *Návrh nekonvenčného zdroja energie na konvenčné účely*. Bakalárska práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2021.

# Impedance spectroscopy of liquid dielectrics

<sup>1</sup>Peter Havran, <sup>2</sup>Roman Cimbala, <sup>3</sup>Miloš Šárpataky

<sup>1</sup> Department of Electric Power Engineering, 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>3</sup> Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical university of Košice, Slovak republic

<sup>1</sup>peter.havran@tuke.sk, <sup>2</sup>roman.cimbala@tuke.sk, <sup>3</sup>milos.sarpataky@tuke.sk

**Abstract** — The publication addresses electrophysical processes in liquid dielectrics based on transformer oils at applied electric field strength, using the method of impedance spectroscopy. The description of the investigated samples and the method of impedance spectroscopy is given in the theoretical part. The practical part clarifies the experiment and analyzes the ongoing electrophysical events in samples of selected liquid insulation materials.

**Keywords** — impedance spectroscopy, complex electrical modulus, complex permittivity, oil

## I. INTRODUCTION

The insulation system must ensure the trouble-free operation of the power equipment at each voltage level. Therefore, it is necessary to examine the condition of the equipment and its insulation system, subject to irreversible physicochemical changes. Transformers are the most important non-rotating devices in the power system, so the need for effective methods of monitoring and examining the condition of the insulation system forces world-renowned scientists to develop and design new materials that improve operational reliability and optimize the operation of all power system equipment [1]. One of the progressive designs is Shell DIALA S4 ZX-1 electrical insulation oil (hereinafter referred to as "SD") from the manufacturer Shell, which is designed to handle the demanding challenges of the latest power transformers. This oil has an extended service life with an emphasis on zero sulphur content thanks to GTL production technology, which converts natural gas into crystal clear base oil [2]. Another electrical insulating oil with very good properties is the inhibited transformer oil MOGUL TRAF0 CZ-A (hereinafter "MO") from the manufacturer Paramo. It is a petroleum product with an oxidation inhibitor that meets domestic and international requirements for the properties of the use of inhibited oils [3]. In this publication, the behaviour of the above-mentioned liquid dielectrics in the applied electric field will be investigated through the impedance method, resp. dielectric spectroscopy in the frequency domain.

## II. IMPEDANCE SPECTROSCOPY IN THE FREQUENCY DOMAIN

The main function of the impedance spectroscopy method is the effect of the electric field on the electric dipoles contained in the dielectric material and the characteristics of their response. Frequency domain impedance spectroscopy has an investigation of the frequency dependence of the complex impedance  $Z^*$ , the complex permittivity  $\epsilon^*$ , and other parameters expressed by them, such as the complex electrical modulus  $M^*$  [4].

Impedance is a more complex parameter than electrical resistance because it involves phase differences and is a basic and major concept in electrical engineering. Impedance spectroscopy is just a "specific branch of the tree" of diagnostic measurements. The size and direction of a planar vector in a right-hand orthogonal axis system can be derived by the vector sum of the components of the real ( $a$ ) and imaginary ( $b$ ) parts:

$$Z = a + jb \quad (1)$$

where  $j$  is an imaginary number with counter clockwise rotation  $\pi/2$  for the  $x$ -axis. Thus, the real part  $a$  is on the real  $x$ -axis and the imaginary part  $b$  is on the imaginary  $y$ -axis. The complex

expression of impedance is also equal to the sum of the real and imaginary part of the complex impedance with the form:

$$Z(\omega) = Z' + jZ'' \quad (2)$$

wherein it has an expression in the plane of Cartesian and polar coordinates, as shown in Fig. 1. Here it is possible to derive two coordinate indicators:

$$Re(Z) = Z' = |Z|. \cos(\theta) \quad (3)$$

$$Im(Z) = Z'' = |Z|. \sin(\theta) \quad (4)$$

where the phase angle and modulus are equal to:

$$\theta = \tan^{-1} \left( \frac{Z''}{Z'} \right) \quad (5)$$

$$|Z| = \sqrt{(Z')^2 + (Z'')^2} \quad (6)$$

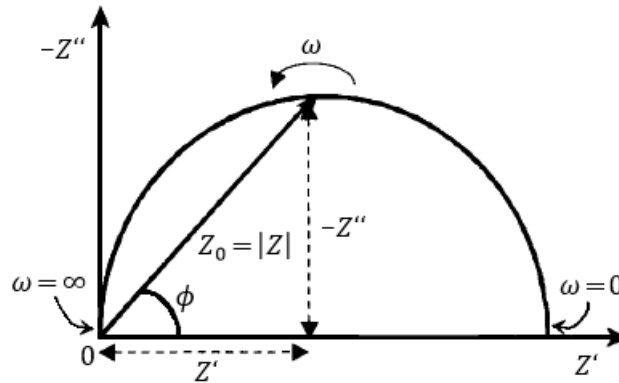


Fig. 1 Impedance vector  $Z$  in a coordinate system [5]

The use of this complex plane is most widespread in mathematics and electrical engineering. In the polar form, the complex impedance is expressed:

$$Z(\omega) = |Z|. e^{j\theta} \quad (7)$$

As is already known, impedance is frequency-dependent. Impedance spectroscopy, therefore, involves measuring impedance over a wide frequency range at an applied electric field strength. This is the resulting response  $Z(\omega)$ , which provides meaningful information about the condition and electrical properties of the insulation material. To interpret the results of the complex impedance  $Z^*$ , it can be very useful to transform the data into the complex permittivity  $\epsilon^*$  or the complex electrical modulus  $M^*$  expressed from it using the relation:

$$M^* = \frac{1}{\epsilon^*} = j\omega C_0 Z^* \quad (8)$$

where  $C_0$  is the geometric capacitance of the electrode arrangement [5].

### III. EXPERIMENT

The measurement of new undamaged electrical insulating oils MO and SD were performed in laboratory conditions at room temperature and atmospheric air pressure. The electrode system Tettex AG Zürich type 2903a was used, which has a volume of 40 ml, an interelectrode distance of 2 mm, and a geometric capacitance of about 60 pF. The frequency range of impedance spectroscopy was captured by an IDAX 300 meter in the frequency band 0.1 mHz - 3 kHz at an applied voltage of 10 V, in other words, a dielectric response at an applied electric field strength of 5 V/mm. The measured data were recorded on a laptop with IDAX 4.1 software. The wiring diagram with a description of the individual instruments is shown in Fig. 2.

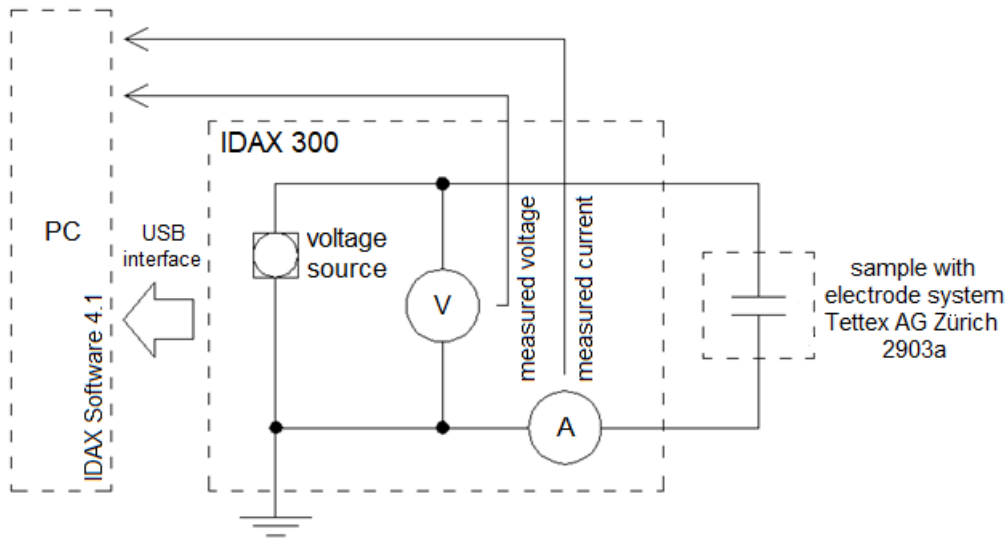


Fig. 2 Wiring diagram of experimental measurement

#### IV. ANALYSIS OF COMPARISON OF LIQUID DIELECTRICS

Frequency impedance spectroscopy can analyze the processes taking place in dielectrics through various parameters. Since the impedance is interchangeable with the permittivity and the electrical modulus, as described by equation (8), this paper will focus on pointing out the processes taking place in the material through these parameters ( $\epsilon^*$  a  $M^*$ ).

Fig. 3 shows a graphical dependence of the comparison of the real part of the complex permittivity of insulating oils on the frequency at an electric field intensity of 5 V/mm. It follows from the above that SD oil has a higher permittivity over the entire frequency range. In the lower frequency range, look for some increase in values caused by the polarization mechanism with some conductivity contributions, especially for SD oil. In the higher frequency band, the permittivity stabilized at 3.2 with SD oil, resp. 2.7 for MO oil. Paraffin oil generally has a relative permittivity of about 2.2. The measured higher permittivity values may be due to imperfect cleaning of the electrode system, which may have resulted in the suspension of microscopic impurities in the oil samples in the form of a higher concentration of molecules per unit volume.

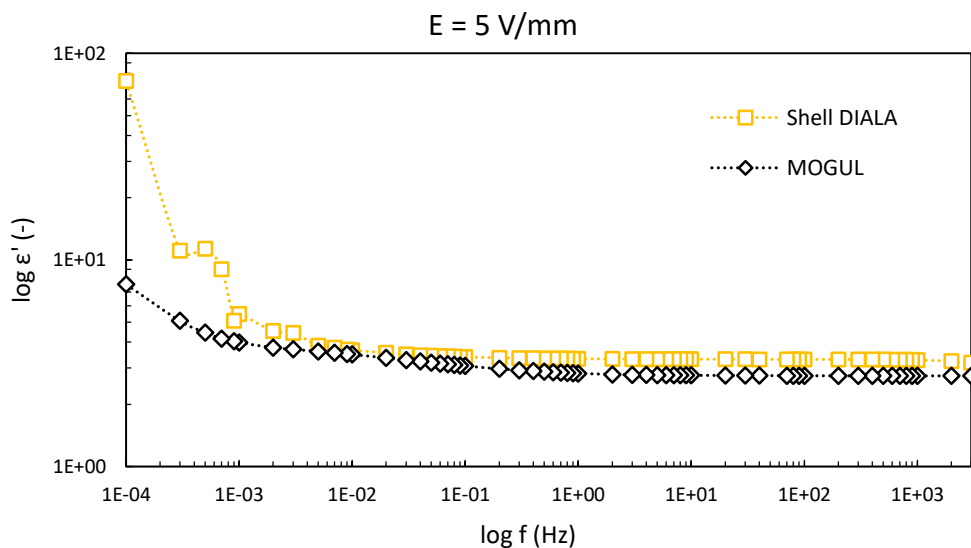


Fig. 3 Comparison of the real part of the complex permittivity of SD and MO oils

The imaginary part of the complex permittivity, characterizing the dielectric losses in Fig. 4, offers interesting characteristics in the whole frequency band, because the comparison of the investigated oils was divided into three frequency bands with different dispersions of dielectric losses. In the low-frequency band, signalling the influence of the conductivity of the material, SD oil dominates, similarly to the real part of the complex permittivity. The mid-frequency region shows higher dielectric losses of MO oil, due to the ripple characteristic, which may

represent a certain polarization mechanism. In the band of higher frequencies, the evolving polarization event of MO oil shifted to higher frequencies. This shift caused a wider capture of the SD oil relaxation process, causing higher dielectric losses in a given frequency spectrum.

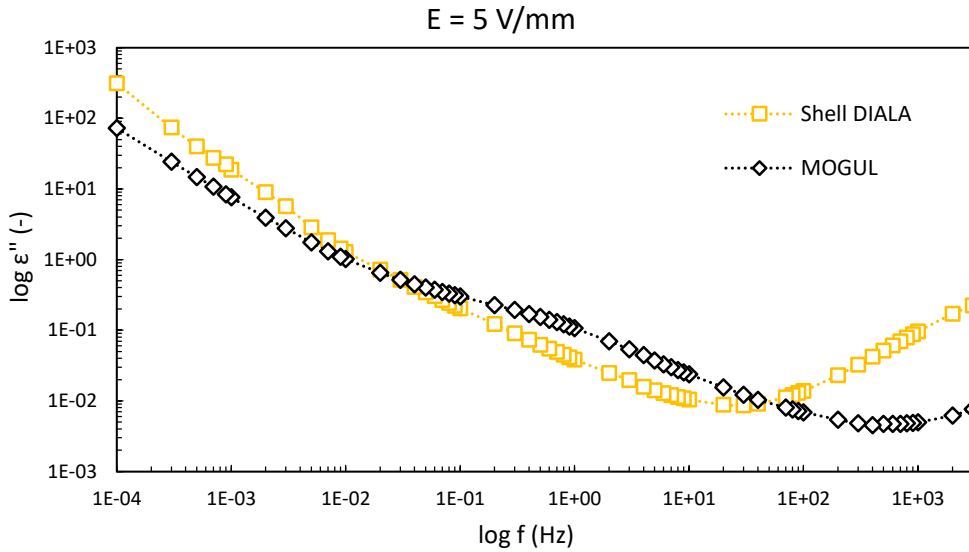


Fig. 4 Comparison of the imaginary part of the complex permittivity of SD and MO oils

By plotting the real and imaginary parts of the complex permittivity in the complex plane of the Cole-Cole graph, it would be necessary to confirm the processes described above in electrical insulating oils. Unfortunately, the complex permittivity in the complex plane does not have meaningful courses to describe the above events, therefore the values of the real and imaginary part of the complex permittivity were recalculated to the values of the real and imaginary part of the complex electrical modulus as a more tellable complex parameter.

Fig. 5 shows the real and imaginary parts of a complex electrical modulus in the frequency spectrum, where several conclusions can be drawn. Since the complex electrical modulus is the inverse of the complex permittivity, all the events described by the complex permittivity will be plotted in the opposite frequency spectrum by the complex electrical modulus. For example, when conductivity processes are present in the low-frequency region with complex permittivity, in a complex electrical modulus they are captured in the higher frequency region, and similarly with polarization events.

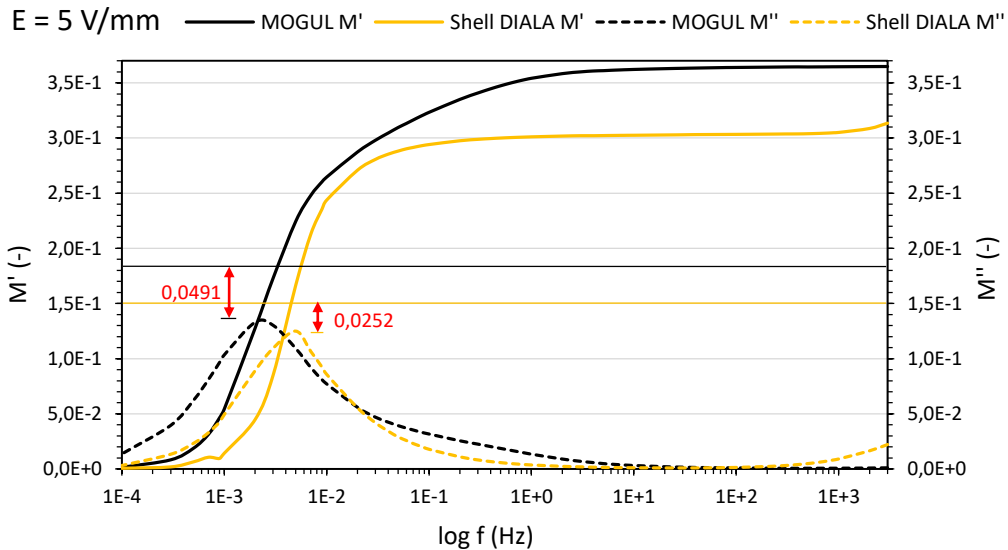


Fig. 5 Comparison of individual components of a complex electrical modulus of SD and MO oils

The graphical dependencies in Fig. 5 show in more detail all the described events which have been presented in Fig. 3 and Fig. 4. With SD oil, the polarization process shifted to higher frequencies, similar to permittivity to lower frequencies. The proof is the hilly curve of  $M''$  and with it the whole curve of  $M'$ . The evolving polarization events in Fig. 4 in the region of higher frequencies are confirmed by hilly curves  $M''$ , while the oil MO had a fine ripple in Fig. 4,

which caused Fig. 5 to widen the waveform to higher frequencies compared to SD oil. This event, in the frequency band from 0.1 Hz to 1 Hz, represents a larger distribution of relaxation times and thus a second captured polarization mechanism in MO oil with approximately 1/3 lower relaxation time. The second polarization mechanism was not captured in SD oil, similar to complex permittivity. However, a slight increase in the characteristics of  $M'$  and  $M''$  was observed in the higher frequency range, which may correspond to the capture of the conductivity process in SD oil. The center lines in Fig. 5 (black and yellow) symbolize the ratio between the real and imaginary parts of a complex electrical modulus using a relationship that represents the ideal state of a dielectric with the only relaxation time that Debye considered:

$$M''_{max} = \frac{1}{2}M'_{max} \quad (9)$$

where it can be seen that SD oil, despite higher permittivity and partially higher dielectric losses, shows a closer Debye behaviour with a difference of 0.0252 than MO oil, which differs from Debye behaviour by 0.0491.

Fig. 6 plots the real and imaginary parts of a complex electrical modulus in the complex Cole-Cole plane, where it is possible to confirm all the above-described events in the frequency spectrum. In the case of SD oil, the presence of a conductivity process in the higher frequency range is confirmed, which manifests itself as a tail behind a quasi-semicircular curve, which expresses an imperfect polarization process. In the case of MO oil, the second polarization event, which has in Fig. 6 fine ripple shape ( $M'$  range 0.3-0.37). The resulting curve of MO oil is thus the sum of two captured polarization processes in the measured frequency band with the distribution of relaxation times.

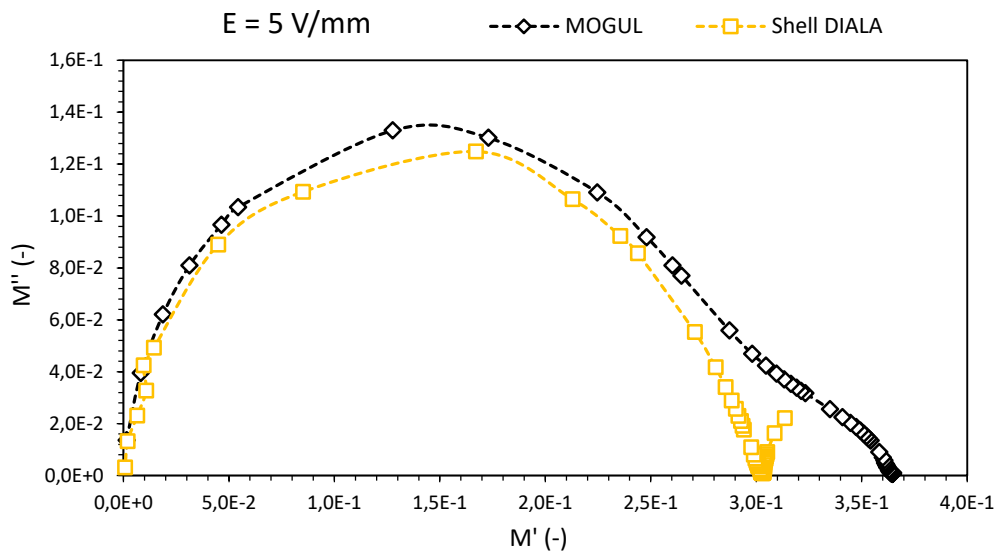


Fig. 6 Cole-Cole diagram of a complex electrical modulus of SD and MO oils

## V. CONCLUSION

This publication aimed to investigate the behaviour of electrical insulating oils SD and MO in the applied electric field by impedance spectroscopy. All the described events through the complex permittivity were confirmed by the complex electrical modulus by its real and imaginary part and also by plotting in the Cole-Cole diagram. From a general point of view, it is not clear which electrical insulating oil shows better dielectric properties at an electric field intensity of 5 V/mm, because two polarizations were recorded in MO oil ( $SD = 1$ ), and on the other hand conductivity process was captured in SD oil ( $MO = 0$ ). These facts will be monitored and confirmed in further research at other applied electric field intensities, to generally describe the behaviour of these liquid insulating materials.

## ACKNOWLEDGMENT

The work was supported by the Ministry of Education, Science, Research and Sport within the project VEGA 2/0011/20 and 1/0154/21 and the Slovak Agency for Research and Development based on contract no. APVV-15-0438, APVV-17-0372, and APVV-18-0160.

## REFERENCES

- [1] P. Havran, R. Cimbala, P. Bartko, „Accelerated Thermal Aging of the Liquid Dielectrics“, 2019 International IEEE Conference and Workshop in Óduba on Electrical and Power Engineering (CANDO-EPE), 207-212 pp. ISBN 978-1-7281-4358-3.

- [2] Shell DIALA Transformers Oils. Online: <<https://www.shell.com/promos/lubes/btb-products/diala-product-range>>.
- [3] PARAMO, MOGUL TRAF0 CZ-A, Technical Data Sheet. Online: <[https://eshop.paramo.cz/data/VyrobkovaDokumentace/ti-trafo\\_cza\\_z5.pdf](https://eshop.paramo.cz/data/VyrobkovaDokumentace/ti-trafo_cza_z5.pdf)>.
- [4] J. R. Macdonald, „Impedance spectroscopy“, In. Annals of Biomedical Engineering. 1991, 20, 2 pp. ISSN 1573-9686.
- [5] E. Barsoukov, J. R. Macdonald, „Impedance spectroscopy: Theory experiment and applications“, New Jersey: John & Sons, Inc. 2005, 608 pp. ISBN 0-471-64749-7.

# Application for Truck Entry and Traffic Tracking

<sup>1</sup>Bohuslav POTOČNÁK, <sup>2</sup>Zdeněk HAVLICE, <sup>3</sup>Lucia VESELOVSKÁ

<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>U.S.Steel Košice, s.r.o, Slovak Republic

<sup>1</sup>bohuslav.potocnak@student.tuke.sk, <sup>2</sup>zdenek.havlice@tuke.sk,

<sup>3</sup>lveselovska@sk.uss.com

**Abstract** – This paper deals with application development for Truck Entry and Traffic Tracking in U.S.Steel Košice, s.r.o. It explains the benefits of using the application during Truck Entry and Traffic Tracking to minimize the time from the entry of the truck to its exit. In addition, this paper is describing the current state of truck inventory in the company. According to the current state, we can analyze and define the requirements of applications to minimize the time.

**Keywords** – traffic tracking, GPS navigation, mobile application, web application

## I. INTRODUCTION

Logistics is a necessary process in the company, which ensures the correct material handling according to time requirements up to the removal of the final product. Every day, a large number of trucks enter the company to unload or load goods. The truck driver has to register manually in the registration room and this registration takes some time. In addition, the truck driver does not know exactly when to enter the area and must constantly monitor the information board. After entering to the plant, the truck driver does not know the exact way to the destination. These problems can be solved by developing mobile and web applications. The mobile application is to be dedicated for truck drivers and the web application for company employees. The tasks of the mobile application are to provide registrations, notify truck drivers, and show the navigation to the destination. The tasks of the web application are to track the current locations of trucks, to see statistics, and to see registrations of truck drivers. Minimizing time would allow raw materials to be stored much faster, and the production shipping process could run faster.

## II. CURRENT STATE OF TRUCK TRANSPORT MANAGEMENT AND LOGISTICS

The current situation of truck transport (see fig. 1) is that the truck arrives at the company and is obliged to register in the registration room. In the registration room, truck driver goes through the registration process and gets an entrance ticket, which entitles the truck to enter the plant. If the registration is not successful, the truck leaves or waits for the entry conditions to be changed. Before entering the plant, the truck is always queued, where, based on occupancy and the destination, it is called to the entrance gate and can only enter the plant. If it is not called to the entrance gate, it must wait in the reserved parking lot to call. Upon entering the premises, the truck is weighed and can continue to the destination where the material handling, such as loading or unloading of goods, is performed, and return to the exit gate, being weighed again before exiting[1]. Weaknesses include the following issues:

- the truck driver does not know the exact route to the destination,
- the truck driver must check the condition and order before he is called to the entrance gate,
- the truck driver does not have an overview of the estimated time of entry into the plant,
- there are insufficient statistics on how long the trucks have been in the area,
- employees do not have an overview of the current movement of trucks in the area,
- employees can only contact the truck by phone.



The disadvantage is the manual registration of the truck at the designated location. Registration could be automated in such a way that an individual truck would be registered to the system through an application. However, registration process is required to be done in the area of company. This could be solved in a way that the truck driver would have to switch to the designated WIFI network upon arrival at the company and would then be prompted by the application for registration. During registration, a unique identifier would be given to the truck for successive interaction with the application.

Truck transport logistics management in the company is shown in the BPMN diagram [2], which is shown below.

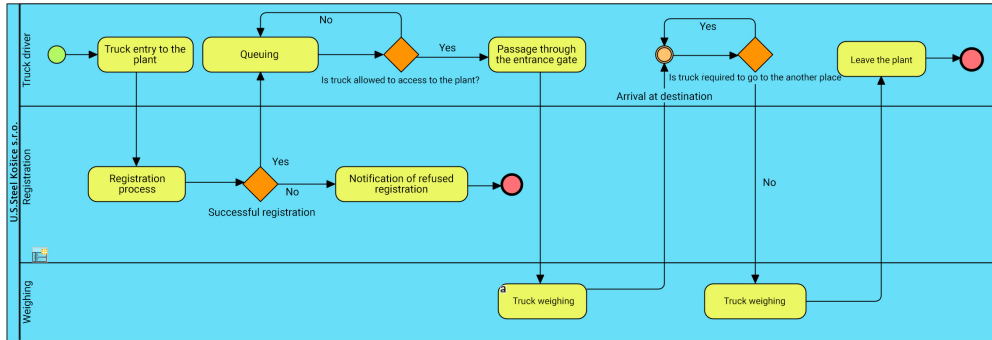


Fig. 1 The current situation of access control and truck transport logistics is shown on the BPMN diagram.

### III. FUNCTIONAL AND NON-FUNCTIONAL REQUIREMENTS OF APPLICATION

Based on the analysis of the current state, requirements and technologies, functional and non-functional requirements for the creation of the application are formulated. Functional and non-functional requirements will be used in the design process and the creation of application models. Finally, these requirements will be evaluated during software validation and verification.

#### A. Functional requirements

Functional requirements define the functionality of the application on the basis of a defined specification and express the behavior of the application. The functional requirements of the application include[1]:

Table 1  
 Functional Requirements

Identification	Requirement
F1.	the truck driver can register via the application
F2.	the truck driver will be informed of the change in the driver's operating status
F3.	the truck will be navigated to the destination and to the place of leaving the plant
F4.	the trucks will be monitored via a web application
F5.	the web application will contain data on the current state of traffic management and logistics
F6.	the web application will contain historical data on the status of traffic management and logistics with the possibility of filtering
F7.	applications will allow communication between truck drivers and the operator

#### B. Non-Functional requirements

Non-functional application requirements define system behavior criteria such as security, reliability, performance, sustainability, portability. The Non-Functional application requirements include[1]:

Table 2  
 Non-Functional Requirements

Identification	Requirement
NF1.	the web application must be run in different browsers
NF2.	the web application must be responsive
NF3.	the mobile application must be run on operating systems iOS and Android themes
NF4.	applications must be extensible
NF5.	applications must be sufficiently powerful

## IV. WEB APPLICATION DEVELOPMENT

### A. Technologies

The application will primarily communicate with the supplied REST API, and therefore the website will be developed only through client-side [3] technologies. The application must be fast, clear and one-page, so that you do not have to wait for the screen to reload during transitions. In this case, it will be selected from three technologies, namely AngularJS, ReactJS and VueJS. The website will be developed through the ReactJS [4] library because we can reuse the source code when creating a mobile application if the mobile application is developed in React Native.

### B. Modules of web application

The web application will contain following modules[1]:

- map module,
- statistic module,
- registration module,
- information and notification module.

The map module shows (see fig. 2) the current location of trucks in the company. Truck information can be displayed by clicking on a specific truck, which is drawn on the map. Truck information is displayed on the right side of the screen. This component can be turned off by clicking on the cross, or it is possible to view a chat with the truck driver. The location of the trucks on the map is updated every 10 seconds without having to reload the page. The statistics module is used for historical browsing of traffic management and truck transport logistics data. The statistics module contains the option of filtering. You can filter by clicking the filter button on the right side of the screen. By filling in the filter data, the filtered data is loaded on the subpage. The registration module is used to display the registrations made by truck drivers. Registrations are displayed for the current day in the time interval from 0:00 to 23:59. The information and notification module contains three basic components, which are:

- events,
- trucks,
- notifications.

The event component is used to summarize the current activities of the trucks that specific trucks perform in the company. The truck component is used to summarize the latest messages written by truck drivers. The Alert component is used to display alerts and restrictions in the company.

The image shown below shows the map module with the current truck position.

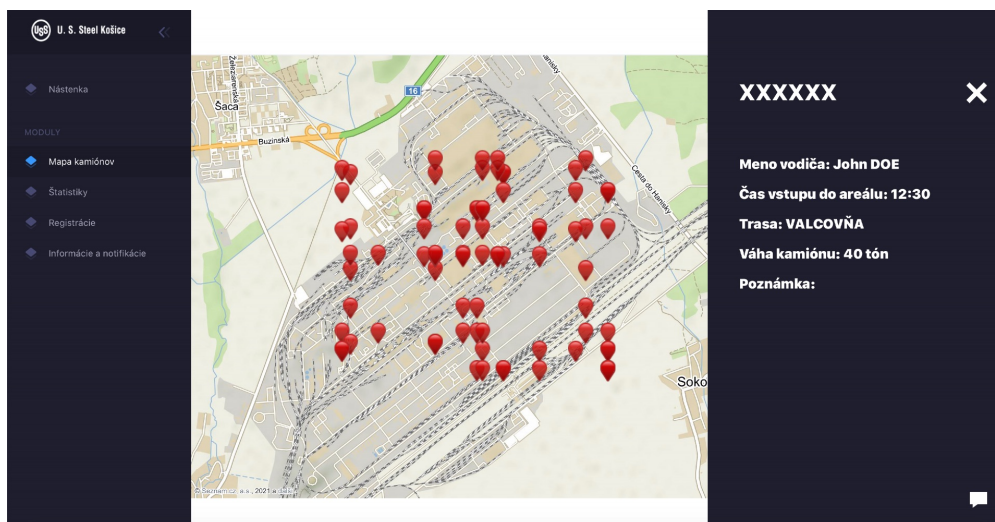


Fig. 2 Current position of trucks on the map

## V. MOBILE APPLICATION DEVELOPMENT

### A. Technologies

The main task of the mobile application will utilize the GPS module with sending the current location to the API. Location services are supported by all technologies described above. The mobile application must also run on both operating systems. Since the web application will

be programmed through the ReactJS library, React Native [5] will be most ideally used when developing a mobile application, as some of the functionality can be downloaded from the web application, the rest will be reprogrammed and a mobile user interface will be created.

### *B. Modules of application*

The mobile application will contain following modules[1]:

- registration module,
- notification module,
- GPS navigation module.

The registration consists of several screens, where it is necessary to fill in:

- personal data, namely the name, surname and number of the identity card or passport,
- vehicle data, namely registration number of the semi-trailer and trailer,
- choice of type, namely loading or unloading of goods,
- and in the case of unloading the goods, it is necessary to enter the order number.

Registration screens also contain screens where you need to approve the terms of privacy. After filling in all the necessary data, it is possible to register. If the registration is successful, the truck driver will receive a registration code. If the registration is unsuccessful, the truck driver must register again.

After obtaining the registration code, the application starts to obtain information to be performed by the truck driver. These activities include:

- entering a reserved parking lot,
- arriving for weighing a truck,
- allow navigation to start,
- unloading or loading of goods,
- and leaving the plant.

When the activity changes, the truck driver is notified by a PUSH notification and a change of the screen in the application.

Once the truck is weighed, the truck driver will be allowed to start navigation. After starting the navigation, a map will be drawn in the mobile application (see fig. 3), which will navigate the truck driver along the specified route. The current position is sent to the company's system in a time interval.

The image shown below shows the GPS module with navigation to the destination for truck driver.

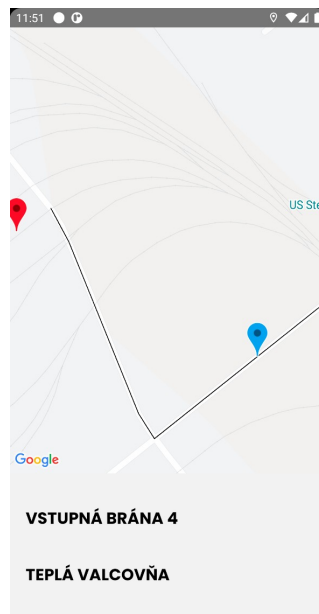


Fig. 3 GPS navigation to the destination

## VI. APPLICATION TESTING

Application testing will be used to verify that the application meets a pre-agreed specification. The application will be tested manually[6], based on unit, integration, system and acceptance testing. Unit testing of the application was already performed during the build of the application, where each component was verified to be working properly. Integration testing of the application was applied if a specific subpage has already been created. Integration testing was used to determine whether the components on a given subpage interacted correctly. An example is integration testing of a map module in a web application. The following checks were performed[1]:

- loading the module itself,
- getting and displaying trucks on the map,
- obtaining additional data about the truck after clicking on the truck,
- updating the position of trucks,
- chat views and messaging options.

## VII. CONCLUSION

System testing was used to check whether the web or mobile application is developed according to a predetermined specification. This testing tested the entire web and mobile application. Errors found during system testing were subsequently corrected and the application was retested. Acceptance testing was performed by the company after previous tests. The web application was tested on the company's dedicated server and the application was tested by two users. The mobile application was tested by designated truck drivers.

The applications were tested in parallel, and the correctness of displaying the trucks position was checked on the web application. In the mobile application, the possibility was checked whether the truck driver could register and whether he was correctly navigated to the destination.

Verification of the web and mobile application will be used to check whether the system meets the functional and non-functional requirements, which were formulated in the analytical part of this work. Verification will be performed on the basis of checking the functionality of the given part of the application with a functional requirement.

The results of the application can be divided into two views, namely the results in terms of security and the results in terms of time. In terms of safety, the implementation of the application has made it possible to monitor trucks in the area. The ability to track trucks ensures this:

- that the truck driver follows the specified route,
- that the truck driver does not stay in the area for longer than necessary.

When driving a truck without the possibility of monitoring the current position of the trucks, it happened that the truck driver stayed in the area longer than necessary. Based on the use of a mobile application, it is possible to monitor the position of trucks, thus eliminating this excess time. Excessive stay in the area results in blocking the entrance by other truck drivers and also a longer time during loading handling at the destination. In terms of time, the time was minimized from the entry of the truck into the area to its departure through the following points:

- Through the application, truck drivers were able to register in an average of two minutes. When registering manually in the registration room, this time was on average 9 minutes. Based on the mobile application, this process was accelerated by 7 minutes.
- Based on the mobile application, the truck driver will be notified when a change is performed. The notification will take place by creating a PUSH notification, so that the truck driver does not have to constantly cut to the information boards and can use this time more efficiently.
- The truck driver is guided to the destination by a precisely defined route, which must be followed. The navigation ensured that the truck driver did not wander around the area and did not stay in places where he should not stay.

Based on the web application, employees can monitor the current state of truck traffic management in the company. A better overview is achieved by tracking trucks on the map and also by the possibility of tracking current statistics, registrations and information. Based on the above results of the application, we can say that the application was made correctly and the objectives of the application were met.

## REFERENCES

- [1] B. Potočník, *Aplikácia pre riadenie vstupu a logistiky kamiónovej dopravy*. Technická univerzita v Košiciach, 2021.
- [2] A. Lynch, "What is bpmn - definition, elements and purpose," 2021, available 02.05.2021. [Online]. Available: <https://www.edrawsoft.com/what-is-bpmn.html>
- [3] G. BREUX, "Client-side vs. server-side vs. pre-rendering for web apps," available 30.04.2021. [Online]. Available: <https://www.toptal.com/front-end/client-side-vs-server-side-pre-rendering>
- [4] W. S. Nahar, "Dynamic view rendering using reactjs and jquery," 2016, available 03.05.2021. [Online]. Available: [https://www.researchgate.net/publication/307144603\\_Dynamic\\_View\\_Rendering\\_Using\\_ReactJS\\_and\\_jQuery](https://www.researchgate.net/publication/307144603_Dynamic_View_Rendering_Using_ReactJS_and_jQuery)
- [5] S. Korolev, "Mobile app development approaches explained," 2019, available 28.04.2021. [Online]. Available: <https://railsware.com/blog/native-vs-hybrid-vs-cross-platform/>
- [6] R. Sing, "Manual testing," 2020, available 02.05.2021. [Online]. Available: <https://www.toolsqa.com/software-testing/manual-testing/>

# System for Processing and Analyzing Test Results

<sup>1</sup>Daniel SLINČÁK, <sup>1</sup>Zdeněk HAVLICE, <sup>2</sup>Richard SZÉPLAKI

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

<sup>2</sup>Siemens Healthineers, Trieda SNP 37, 04011 Košice, Slovak Republic

<sup>1</sup>daniel.slincak@student.tuke.sk, <sup>1</sup>zdenek.havlice@tuke.sk, <sup>2</sup>richard.szeplaki@siemens-healthineers.com

## Abstract

**Abstract** – This paper presents a system for handling and analysing test results. The system was created with cooperation of the company Siemens Healthineers. This system is in the form of a web application, where users will be able to interactively view and analyze results from the testing process.

## Index Terms

**Keywords** – test analysis, test results databases, web application, testing

## I. INTRODUCTION

Testing is an important phase in the development life cycle of each software and hardware. The customer expects a complete and functional product. In some types of products, for example in medical devices, testing is an mandatory and regulated part of the development process. The main reason is that it could affect the health and life of patients, therefore the testing process is complex and long.

This leads to a big number of test products, called artifacts [6], such as test runs, test suites, individual test cases and their results. Managing all these products could take a lot of time, so the goal is having them saved in a practical way, so it is possible to easily analyze them. Processing these products, comparing them with new and historical data and finding relationships between them will help with the testing process and the development itself, both in present and future projects.

## II. GOALS

The goal of this work is to design, implement, test and deploy a software system that will help with processing and analyzing current and historical test products. The system will also provide an overview of the whole testing environment, which includes all the test machines and their components.

This system will be reusable, robust, secure, maintainable and scalable, to provide the option to easily add new features and use it on different types of projects, where the analysis of test results is important.

## III. SOLUTION

All the features and traits of the final system, described in the previous section Goals, were derived by gathering requirements from future users and stakeholders of the system. By analyzing these requirements, we were then able to design and review how the system should be created, how it should behave and look. This process is described in the published thesis in [3]. In this section we focus on some important aspects and parts of the system, which helped to satisfy these goals and requirements.

### A. Iterative development and testing

The system was developed using iterative development methods. Thanks to this type of development, users could also use the latest stable version while new features were being developed.

Whenever a new feature for the system was completed, a new version of the system was deployed to be tested by users. Getting the feedback from the future users as soon as possible was important and helped to deliver a useful and usable system, that would help them in their work.

### *B. MVC Architecture*

The system uses the Model-View-Controller architecture [5]. This architecture splits the system logic into three types of components - Models, Views and Controllers, where each one has a specific functionality.

The Model component represents the data, which should be displayed on a given page in the system. The structure and look of the page is defined in the View component. Every interaction and requests from the users go through the Controller component.

Having the system be designed with the MVC architecture from the start is important in many aspects, such as allowing to implement and test each component independent from the others. Future improvements and modifications of the system will be also easier to plan and develop, since just having these types of components helps to understand where the new features should be implemented.

The system often uses a ViewModel component instead of just Model. The difference is that the Model component represents one type of data, in most cases a single record from a database table. A ViewModel can include one or more records and also additional data. This is also helpful when thinking about the future of the system, when new requirements could and probably will be introduced.

### *C. ORM for working with databases*

An important part of the final system are the databases it uses. These databases were designed by analyzing the test artifacts that will be created in the testing process. After that it was possible to create diagrams explaining relationships between these artifacts, to better visualise how the final databases should look.

After these diagrams were finished and we had a good idea on how the design of the databases, the only thing remaining was to create them. The method used for creating the databases, their tables, attributes and relationships Object-Relational Mapping (ORM) [1] using the framework called Entity Framework (EF) [2]. With ORM, the databases are created from object-oriented classes, specifically in C# for EF. For each database table a class was created, where its public attributes represented the attributes for that table and references to other classes would be translated as relationships to other tables. EF was then used in the whole system to send queries to the databases and receive the desired data Using ORM with EF had several advantages:

- 1) No SQL needed to be written to create the databases and their entities
- 2) Easier work with making queries on the databases
- 3) Query results were saved to object-oriented classes
- 4) Basic SQL injection protection when accepting user inputs
- 5) Databases can be maintained and modified both from the system using EF and other database management systems

### *D. Making the system secure*

The security of the system is extremely important since the system handles loads of data from the testing process and the testing environment. Unauthorized users should not have the ability to view this data and not every authorized user will have the same permissions in the system. The two processes that help to secure the system are authentication and authorization. Both of these methods are applied in Controller and View components of the MVC architecture.

Authentication is used to verify the identity of the user, which is trying to get access to the system, which is achieved by using individual user accounts. When an unauthenticated user wants to display a page, he will be instead redirected to the login page.

Authorization is a process, in which the system checks if the user has the required permission to perform an action. This is done by checking the role of the user and the defined policy for some actions. There are currently 4 types of roles defined in the system: Administrator, DevOps, TestOps and Basic. Policies are defined for each action that should be authorized. For example, to have the ability to create a new record for a test machine, the user must satisfy the ModifyTestEnvironment policy. This policy specifies that the user must be either in the Administrator or DevOps role.

If a user does not satisfy the required policy, the page or some part of the page will not be displayed to him. But security by obscurity is not enough, since some malicious users could find the URL address for pages that send POST requests to the server. To protect against this, authorization is also used when users are requesting the desired pages. Unauthorized users will get an Access denied message when trying to access these sites.

## IV. BENEFITS PROVIDED BY USING THE SYSTEM

The final system and its usage can help individual users, teams and the company in multiple ways. Detailed description on how to use the system, how it can be used and other details are described in the published thesis [3]. In this section we will focus on three main use cases - managing the test environment, processing and analysing test results and accessing the raw data from the system.

### *A. Management of the test environment*

The system provides the option to easily manage, modify and export information about the whole test environment. Users can see each available test machine that is being used to execute tests. Individual hardware components and software versions on each test machine are also available for users to view. These include:

- 1) Graphics cards
- 2) Processors
- 3) Motherboards
- 4) Physical memory
- 5) Logical memory
- 6) Network configuration

Users, which are authorized as described in the section 3.3, can also modify the information about the test machines and their components, directly in the system. It is possible to add a new test machine to the test environment and modify or delete an existing one. The same applies to each component of the machines.

Having the information about the test environment and also the tests and their results in the same place is really beneficial for every person involved in the testing process. The outcome and duration of one test can differ on different test machines. Being able to see the historical difference between the execution times and results on the same machine or between different machines will help the testers and developers to make the testing process more accurate and therefore also improve the tested software itself.

### *B. Viewing and navigating the test artifacts*

The main purpose of the system is to process and help to analyse test results, which can be in the magnitude of thousands in a single day. The system provides a user-friendly way to view these results and an easy way to navigate between them and other test artifacts [6]. Users can view:

- 1) Test Cases
- 2) Test Suites
- 3) Test Runs
- 4) Identified defects in the tested software
- 5) Reports from testing

Not every piece of information about these artifacts is important to every user of the system, that is why only the most important are displayed first. Users that will need to perform a thorough analysis can then find all the necessary information in the detailed view page. By also having the links to other test artifacts, they can easily navigate from the low level artifacts to the high level during their analysis.

There are two ways how these artifacts can be displayed on a given page:

- 1) Grid view - items are displayed in a grid, where each item is in the form of a tile or card, which displays the most important information. An example of the grid view for displaying test runs is displayed in the figure 1.
- 2) Table view - items are displayed in a table, where each item is in a separate row. The table also displays only the most important information about the displayed items.

Users can toggle between these two views on every page. They can also modify how many records of the test artifacts they want to have displayed on a single page by applying a variety of filters.

When selecting a specific test artifact, users will be able to see a detailed view with all the available information about it, charts with useful visualisations and direct links to all other associated artifacts. For example, when viewing all test runs, only the most important information is available - how many test suites does it contains, for which version of the tested software it was used, what is the final outcome and others. When viewing a specific test run, the detailed page contains all the available data and references to all the test suites, which can be also displayed in detail.

### *C. Analysing charts*

Like it was mentioned in the previous subsection, a lot of pages contain charts with additional information about the test artifacts. Charts on each page provide a useful visualisation of the available data for users, which helps to reduce the time and effort needed to perform the complete analysis.

These charts are interactive, users can hover and interact with the displayed data in them. The colours in the charts are selected to be as easy as possible to view, since most of the time, a single chart contains a lot of items.

In the current version of the system, which was described in the diploma thesis [3], charts are predefined - that means, that users cannot change the type and data in them. All of the charts, which are currently present in the system, were selected by analysing the test artifacts and by having dialogs with the intended users of the system. Next improvements for the charts would be to allow users to create them according to their specific needs.



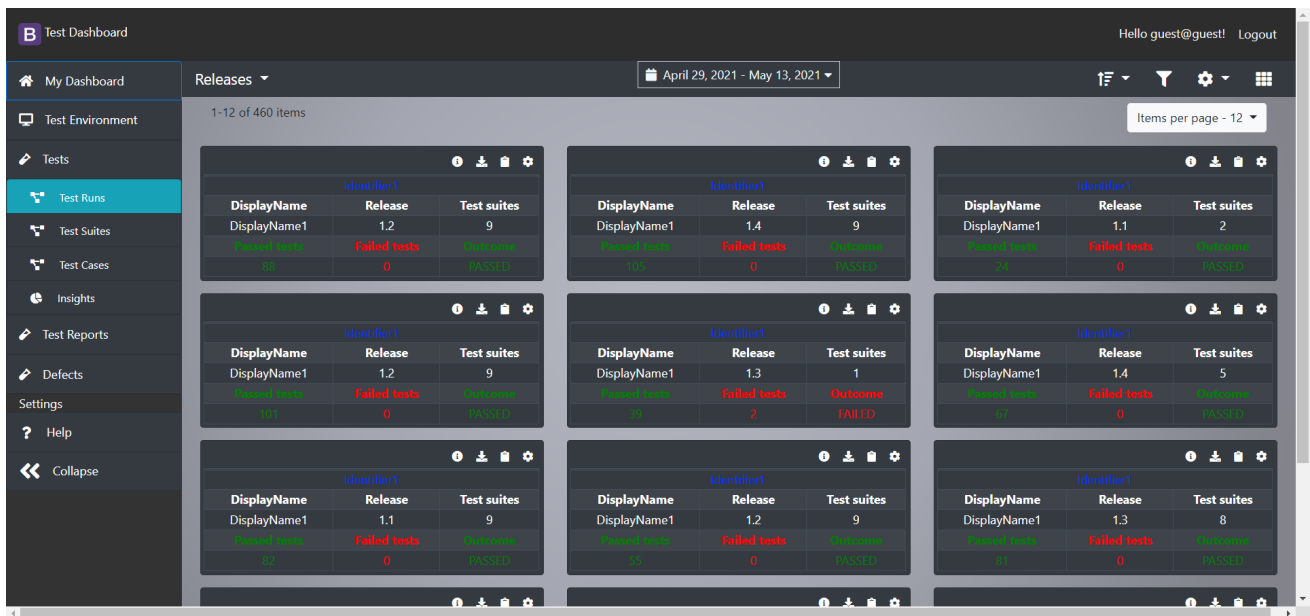


Fig. 1. Viewing test artifacts in a grid view

#### D. Tracking identified defects

During the analysis of the test artifacts, users can find some defects or other error types in the tested software. These can be then recorded directly in the system or in the internal defect management tool the company uses.

Only authorized users can create these new records for identified defects. There are two ways on how to create a new record:

- 1) Manually filling out all of the required information about defect management.
- 2) Linking to an existing record in the internal defect management tool, which will then automatically create a new record in the system with all the available information.

All of the identified defects can be easily tracked in the system, as it is possible to view defects currently present in the software and all already resolved defects. This also provides some new ways on how to improve the complete analysis of the testing process - it is now possible to combine data available about testing machines, tests and identified defects in the software. So for example, users could start to see a pattern, when a particular type of defect is present only on one test machine.

#### E. Sharing testing reports

Detailed testing reports are important for every user which has a part in the testing or even developing the software. These reports contain an overview on the whole state of the tested software and links to all other test artifacts that were used in creating the report.

Multiple versions or branches of the system can be tested independently, so the system provides an easy way to view the latest report from each one. Like all other test artifacts present in the system, users they list all reports and search or filter them according to their parameters.

These reports can be shared by printing the report to a PDF file or by just sharing the report page of the system.

#### F. Accessing raw data

Displaying the data in the system for users in an easily readable way is really important, but sometimes its not enough. What if some users want to somehow process the data in some other external tools? They would have to manually copy and enter the data. But what if they want to automate it? Parsing the HTML file of the displayed page is one solution, but its a terrible one.

The correct solution is to offer users the ability to get their desired data in a uniform way. This is accomplished by having a REST API[4], which can be described as an interface where users send their request to the system and receive back the data.

For example, if a user wants to display the test cases executed in the last month, he would navigate to the test cases page and he would filter the results with the calendar control. But if he wants to get the data in a structured way, he would send a request to the REST API with the desired date range and receive a JSON with a list of the test cases. The user can then use the JSON and its content in any way he desires.

The system supports REST API requests for multiple different test artifacts with several options on how to modify the query. A detailed documentation of all available requests, parameters and example outputs is available to view directly through the system.

## V. EVALUATION ON A PROJECT

The system was integrated into the project for developing medical devices, where thousands of tests are executed each day. The system was used to store information about the testing environment and several test artifacts from the testing process, such as test runs, test suites, test cases and identified defects in the system.

TestOps and DevOps users in the project were using the system for 4 days to help them in their daily activities. This included getting the latest data about the available test machines in the test environment and then executing all the tests.

The test results and other artifacts were saved to the databases. These databases were also used to check if all required tests were successfully executed and executing them again if something went wrong.

The figure 2 shows the total duration of the testing process before and after integrating the system into the project. Four measurements were made, where the average duration before using the system was 439 minutes and 353 minutes when the system was a part of the testing process. The system helped to speed up the process, where we observed an average of 19.5% reduction of the total duration. This is a significant improvement with multiple benefits for the whole team, since the execution of all the tests and their analysis are done on a daily basis. The saved time can be then used to improve the existing tests and testing processes, or create new ones.

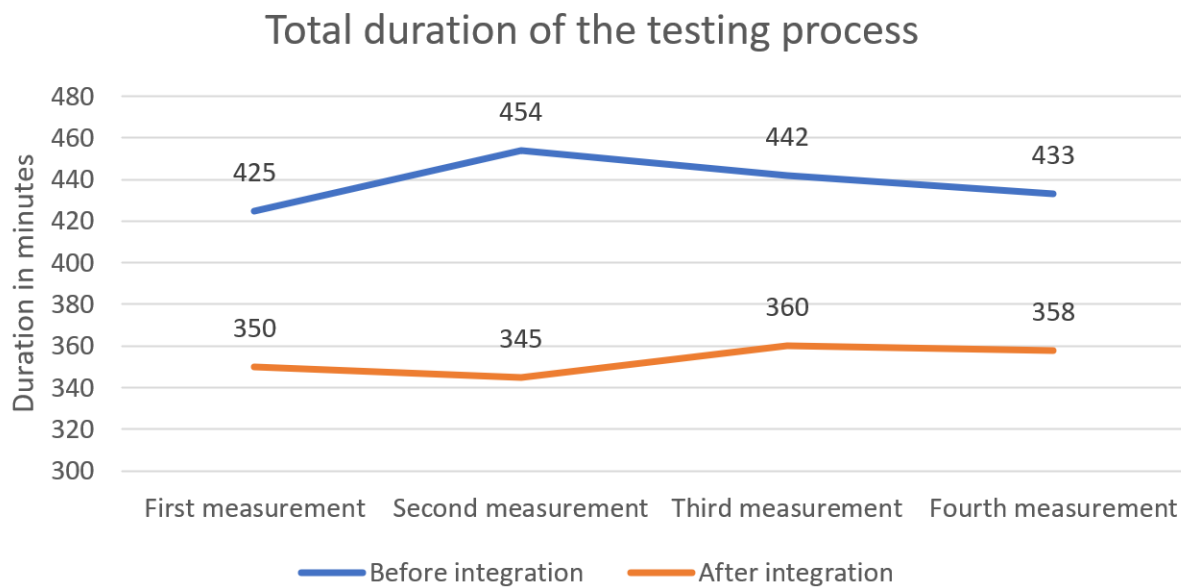


Fig. 2. Duration of the testing process

## VI. USER FEEDBACK

Dialogs and conversations with users were used throughout the whole life cycle of the system. After the deployment and evaluation of the current version system, as described in the diploma thesis [3], we gathered some final thoughts from the users which were actively using the system.

The asked users had a very positive feedback on using the system. The main positive comments, in no particular order:

- 1) Useful charts for each test artifact in the system
- 2) Easy navigation between all test artifacts
- 3) Detailed information about each test machine
- 4) Easy to understand user interface of the whole system
- 5) Being able to use the system on computers, tablets or phones
- 6) Supporting a REST API for getting the data in a structured way
- 7) Performance of the system

Users also had some constructive ideas on how to improve the system, by either changing what information is available in some pages or what new features could be developed to make the system more useful. Some of these ideas are:

- 1) Ability to create own charts
- 2) Comparing two of the same test artifacts
- 3) Having more complex filter options
- 4) Exporting displayed data in multiple formats

### 5) Creating automated notifications

Thanks to all of these responses, we can say that the system and its usage were a success. By also having the constructive comments, we can plan on how to make the even more useful.

## VII. CONCLUSION

The goal of this work was to design, implement, test and deploy a system for processing and analysing test results. The final system was integrated into the testing process in one project, where it was used to help with the testing of medical devices. Traceability from the highest level of test artifacts to the lowest level helped with the analysis of test results. Users could also work with detailed information and view useful charts for each test artifacts. They could compare and generate new test reports and record detected defects in the tested software. The system also gave users the ability to view and edit information about the testing environment, which includes all the testing machines and their components.

The testing process was improved in several ways thanks to this work, where the total duration was reduced by a significant amount.

The users were satisfied with the developed system. By having continuous dialogs with the users about how they were using the system, new improvements and features were identified, which will be useful to further improve the usefulness of the system.

## REFERENCES

- [1] SCHWICHTENBERG, Holger. *Modern Data Access with Entity Framework Core*. Apress, Essen, Germany, 2018. ISBN 978-1-4842-3552-2.
- [2] ROULEAU, Derek J. *Beginning Entity Framework Core 2.0*. Apress, East Baldwin, Maine, USA, 2019. ISBN: 978-1-4842-3375-7.
- [3] SLINČÁK, Daniel. *System for Processing and Analyzing Test Results*. Diploma thesis, DCI FEEI Technical University of Košice, 2021.
- [4] REYNDERS, Fanie. *Modern API Design with ASP.NET Core 2*. Apress, Odijk, The Netherlands, 2018. ISBN: 978-1-4842-3519-5.
- [5] DEACON, John. *Model-view-controller (mvc) architecture*. [Online]. 2009. [viewed 2021-01-24] <http://www.jdl.co.uk/briefings/MVC.pdf>
- [6] PARIZI, Reza Meimandi; LEE, Sai Peck; DABBAGH, Mohammad. Achievements and challenges in state-of-the-art software traceability between test and code artifacts. *IEEE Transactions on Reliability*, 2014, 63.4: 913-926.

# Doc2vec as an expansion to full-text search engine for a specific domain

<sup>1</sup>Jerguš KAČMÁR, <sup>2</sup>William STEINGARTNER

<sup>1,2</sup>Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice  
Letná 9, 04200 Košice, Slovak Republic

<sup>1</sup>jergus.kacmar@student.tuke.sk, <sup>2</sup>william.steingartner@tuke.sk

**Abstract** – In this paper, we tested a simplistic solution to improve the relevancy of results given by a full-text search engine. The need for this improvement arose on a system that searches for similar documents to a one given by a user or to a phrase used for searching. The previous system used full-text search capabilities of Elasticsearch and we added and tested Doc2Vec module that serves as a semantic search engine. We needed to train our own NLP model because we were working with a specific technical domain and a pre-trained public model wouldn't cover all the specific terms and wouldn't capture nuanced meanings of words specific to our domain.

**Keywords** – Doc2Vec, Search engine, NLP, Word2Vec

## I. INTRODUCTION

When we have a set of documents in which we need to have the ability to look for a specific one, usually the quickest solution is to use lexical matching of characters. While such a method is powerful and often sufficient, there is still room for improvement. Let's say we can utilize existing methods of natural language processing to give a user the ability to search not just with the exact phrase, but with the idea of the phrase as well. With this capability, we can provide him with results that are close in meaning to what he entered. This has multiple benefits such as returning documents that may not necessarily contain specific words that are in a query phrase, but which have the same or similar meaning to that phrase. Such documents could remain hidden from a user in a system that needs a strictly exact form of words and phrases.

After we find a suitable solution to this cause we will try to combine it with the already existing ones so that users that are already comfortable with the tool, and won't lose any previous functionality.

The paper is organized as follows. In Section II, we discuss the motivation for our work and we present an actual problem that is a topic of this paper. Then, in Section III, an analysis of our solution is provided. Section IV presents in details the most important aspects about the implementation of our solution. The survey of results and findings is given in Section V. Finally, the Section VI concludes our paper.

## II. DEFINING PROBLEM

In this section, we will describe the domain we are working with, a tool that we want to enhance and what we want to accomplish with this enhancement.

### A. Current solution

The tool in question is a result of the project and research of Vladimír Kovalský [1]. It is providing a search of templates for tickets on a system where users report changes on internal systems in an organization providing ICT solutions. Users create report tickets and when the report has the potential to be repeated, a template is made from that ticket. But specific information isn't included in that template. Instead only generic descriptions are left in templates for another user to expand upon when creating his report. Thus, when we want to create a search engine for those templates, we search inside text fields of historical tickets instead of templates because templates contain only a small amount of information that is often insufficient for our needs. When a match

is found with data inside a ticket, a template from which the ticket was made is then shown to the user.

All the text fields of tickets are stored inside Elasticsearch<sup>1</sup> indexes with additional information about whether the ticket was made from a template and identifier of that template if so. For extraction of semantic meaning, we can only use title and description of tickets, because those fields contain actual sentences in which we can find coherent ideas.

Search is done with Elasticsearch queries that utilize TF/IDF<sup>2</sup> algorithm. Results are sorted with custom Elasticsearch metric that gives us information about how relevant are results to a query, but we can only rely on that metric for sorting of results since we can't tell whether the result is relevant as that is subjective and we can only provide the user with our best guess. It is then problematic to make this metric interact with the metric that we will use for NLP when we want to have both functionalities combined in one place.

### *B. Argument for NLP*

With technologies advancing the search engines we use regularly are utilising NLP to give us the most relevant results to our queries. With advancements in technologies, user's expectations are matching the trend and to satisfy them we need to produce more powerful tools while keeping them easy to use. One way to make a tool such as a search engine easy to use is to recommend to a user the results that we think are suitable despite them not containing specified phrase in situations when the exact phrase wasn't found. We want to avoid situations when a user is presented with no results because for his phrase there isn't an exact match while rephrasing the query would lead to getting the desired result. The user may not try multiple times till he thinks of the correct query and may instead conclude that we don't have his result in our database. An experienced user is more likely to make a correct query but an inexperienced one can benefit greatly from just giving us his idea without worrying about what is the most likely phrase to match his desired result.

There are also things like recovering from incorrect spelling or simply the fact that we can't tell that the result is relevant simply because it contains part of the query phrase. With the ability to extract meaning from words we can then have a metric that will tell us how distant are query and result in meaning.

## III. PROPOSED SOLUTION

In this section, we explore why we choose Doc2Vec and how used it to improve our search engine.

### *A. Tasks*

We are working purely with one input field on the side of a user. He can input random text or identification of document that we have in the database. We need to return a list of documents with their semantic distance to the inputted text.

Let's consider the user's input as another document. We need to have a method that will tell us the distance between two documents. For that, we need to place documents into space. This space will be defined by vectors and all the documents contained in it must have their vector. This process is often referred to as the embedding of words or documents into space. When we have that then the calculating distance between two documents becomes straightforward.

We return documents with the shortest distance to the one input to the frontend side and display them to a user as documents with the most similar semantic meaning.

### *B. Choosing of Doc2Vec*

Doc2Vec method does exactly what we need. It assigns vectors to documents we feed it. It's not a brand new method as it was introduced in 2014[2], but it is a proven method that's still being used today. There are multiple approaches to calculating vectors for words or documents. Some other older competitors of Doc2Vec such as GloVe<sup>3</sup> yield almost identical results[3]. Other newer approaches are too robust for our use case, such as Bert<sup>4</sup> which offers a multitude of other functionalities we wouldn't use and would require us to modify the model to even begin to approach the speed of Doc2Vec[4]. With these vector spaces of words and documents, the biggest issue for us was the fact that adding new words and vectors without recalculation the position

<sup>1</sup>Search engine <https://www.elastic.co/>

<sup>2</sup>Term frequency-inverse document frequency is a numerical statistic used in search engines.

<sup>3</sup>Global Vectors is an algorithm for embedding words into vector space.

<sup>4</sup>Bidirectional Encoder Representations from Transformers is an NLP model developed by Google.

of all the others is lowering the overall accuracy of the model. Usually, with this approach to searching, we need to periodically recalculate the whole vector space because all elements in it are affecting each other. This task of calculating vector space takes time and it takes a lot more time when we choose a robust model.

In the end, Doc2Vec is one of the least computationally difficult models to create while still giving impressive results. We used the Skip-gram method which extracts vectors for words and documents from a simple neural network, which is trained to predict the context of a given word[5]. This process was streamlined by using the Gensim library which makes this easy even for beginners in the NLP field.

#### IV. IMPLEMENTATION

In this section, we describe the steps taken to implement our solution.

##### A. Training NLP model

To get our desired vector space, we used the library Gensim. We need to provide tagged documents to the train function. Tagged documents are simply text strings with identifiers. To get our final string of text, we took all text fields of report tickets that contain coherent sentences and concatenated them. And for identifiers, we used IDs of tickets so that we can easily link them back to the original dataset.

##### Generization

All concrete data like dates and codes aren't fit to be left in our dataset for us to be calculating vectors for them. Through a set of regular expression checks, they are replaced with generic words like "date" and "code". If a user wants to find these specific data, we will rely on a full-text search.

##### Tokenization

We want our final data to be segmented into words with clear delimitations. For that, we used library spaCy<sup>5</sup>. And to try to turn those words into their base forms, we used again the same library with its lemmatisation function that also takes a parameter of what language we are working with. To guess in which language the document is we used PYCLD2 library<sup>6</sup> because our data are multilingual.

##### Training

Training our model is simply calling Doc2Vec function. We fed it our data and used the following parameters.

Training parameters	
Parameter	Value
Dimensionality of vectors	vector_size=300
Distributed memory	dm=0
Hierarchical softmax	hs=0
Negative sampling	negative=5
Initial learning rate	alpha=0.01
Final learning rate	min_alpha=0.0001
Minimum frequency of word	min_count=1
Context window	window=15
Downsampling	sample=1e-5
Iterations over corpus	epochs=10
Generation of word vectors	dbow_words=1

Table 1  
Table of parameters specified for training of Doc2Vec model.

Since training doesn't finish in real-time, we can't retrain a new model immediately when new data arrive and instead, we set automatic retraining of the model to once every week on a designated machine with newly added tickets. Automatic scheduling was set up with Celery framework<sup>7</sup>.

<sup>5</sup>spaCy is a library for Python language used for various NLP tasks

<sup>6</sup>Python Compact Language Detect 2

<sup>7</sup>Celery is a task queue framework in Python language with scheduling capabilities.

### *B. Combining NLP search with full-text search*

We couldn't find a meaningful interaction between the metric from Elasticsearch and our semantic similarity. We still wanted to combine both searches into one tool, so we tried to set up our own rules for when is NLP result preferred over the lexical one and the other way around.

#### *Semantic similarity*

If the user inputs identifier of a document then we should have its vector already calculated and just fetch it. If the user inputs random free text, then we have to infer its vector. For that, we use Gensim's function for inferring vectors `infer_vector` that tries to place this vector somewhere into vector space based on words that it contains. Words that weren't present at the time of training are discarded.

To measure the distance between two documents, we choose cosine distance, which equals one if two vectors in question are pointing in the same direction and are zero if vectors are pointing in opposite direction. This multiplied by one hundred is presented to a user as a percentage of similarity.

#### *Rules for sorting of results*

Since these two measurement systems don't mix well, we set a hard cutoff for when is NLP result preferred over the full-text one. A suitable cutoff will vary for different models, our was set to 0.70 or as the user sees it 70%. The reasoning is that above this threshold a match will very likely be related to what we are searching for. After that, we placed full-text results that contain whole or part of a search query in their fields. Below that we put leftover NLP results that are in the range 70% to 50%. Below 50% they likely won't be of any use. Results that are found by both methods are elevated to the top. If we get results that are found with both methods, they are very likely interesting to the user, we can then safely assume that result is truly relevant.

## V. FINDINGS

Looking over examples of related words, we were a little sceptical about the efficacy of our model. We found that we had a lot of data that wasn't useful for word embedding. Lots of repeated data, samples in different languages, descriptions too short, not enough variety in used vocabulary meant that our at first vast dataset was barely sufficient for this task. Examples of words evaluated as related by our model are shown in table 2. Many times we find word relations that are natural to human, but it's not as impressive as publicly available models for general use.

Word embeddings are giving us an idea of how the model will perform, but not the whole picture. The task was to find similar tickets through a free-text description. Finding documents seems to be sufficient for users. Related documents are indeed clustered together in vector space, and searching with a short description of a ticket is giving us results mentioning the same thing even if it isn't worded exactly alike.

We gathered nine potential users that are knowledgeable in our domain and had them test our tool. After testing they filled a questionnaire asking them about their opinion on the relevancy of results. All of the testers deemed the expansion of the previous tool as a welcome improvement. In a question about the relevancy of results, we were met with success for the most part while two of the users had problems finding results they were searching for which was later attributed to our database missing the data in question. Results were marked by their source (either Elasticsearch or Doc2Vec) and ended up having fairly even representation. Positive feedback let us deem this approach effective for this task.

One cause for concern is that searching through a phrase will favour documents that have short description because the additional text makes the document deviate its overall meaning away from the idea of query phrase. The opposite effect can be seen with full-text searching which has a higher chance of finding matching words in longer text. Thankfully we can often time find the same documents with both approaches and their elevation to the top of results is then hopefully giving a user the best match. If only the Doc2Vec approach would be used then documents with long texts may be harder to find than others.

## VI. CONCLUSION

We set out to improve the relevancy of results on a full-text search tool for finding similar documents. We choose to employ NLP technologies specifically Doc2Vec because of its ease of use with the use of finished libraries and its ratio of performance to the time it takes to train such a model. The model had to be trained anew by us because no general use model would cover our domain-specific terminology. We found that our dataset wasn't ideal for this task, but it produced

Searched word	Closest words and their cosine distance to searched word				
firewall	0.8492 fw	0.7567 vdom	0.6798 rule	0.5879 permission	0.5614 open
install	0.6957 new	0.5494 configure	0.5222 subscription	0.5165 non	0.5122 server
server	0.6013 application	0.5680 computer	0.5267 systemid	0.5198 prod	0.5122 install
email	0.6913 contact	0.5354 phone	0.5337 emailaddress	0.4945 hotline	0.4633 attachment
time	0.7433 hour	0.6770 lead	0.5979 work	0.5920 day	0.5164 emergency
question	0.8204 answer	0.7478 fit	0.7326 credo	0.7053 note	0.5475 prerequisite
system	0.5808 os	0.5580 responsible	0.5554 inform	0.5299 hostname	0.4831 downtime
document	0.6523 change	0.6398 mandatory	0.5986 accord	0.5869 attach	0.5862 implementation
request	0.6234 requestor	0.6106 description	0.5977 add	0.5798 use	0.5717 change
update	0.5694 stop	0.5342 patch	0.5098 version	0.4955 activity	0.4775 vernumber
level	0.7206 component	0.6795 something	0.6767 stable	0.6693 outage	0.6601 ready

Table 2  
 Table of examples of related words search with Word2Vec

sufficient results for us to believe that we improved the relevancy of results in many cases. We combined the search capabilities of Elasticsearch with Doc2Vec search of similar documents and elevated those results that were found by both methods to the top of the list. We automated training of a new model for when new documents are to be added to our dataset. We received positive feedback on the performance of the new search method and finished a tool that is ready for deployment in a production environment.

#### ACKNOWLEDGMENT

The first author is grateful to a colleague Ing. Vladimír Banóci, PhD., for coming up with the original idea for this project and giving me access to his previous NLP experiments in the domain of our project.

This publication was realized with support of the Operational Programme Integrated Infrastructure in frame of the project: Intelligent systems for UAV real-time operation and data processing, code ITMS2014+: 313011V422 and co-financed by the European Regional Development Fund.

#### REFERENCES

- [1] V. Kovalský, "Search for models for change management in computer systems," Technical University of Košice, Košice, Tech. Rep., 2020, (in Slovak).
- [2] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proceedings of the 31st International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, E. P. Xing and T. Jebara, Eds., vol. 32, no. 2. Beijing, China: PMLR, 22–24 Jun 2014, pp. 1188–1196.
- [3] T. Lin. (2010, Jul.) Performance of different word embeddings on text classification. [Online]. Available: <https://towardsdatascience.com/nlp-performance-of-different-word-embeddings-on-text-classification-de648c6262b>
- [4] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019. [Online]. Available: <https://arxiv.org/abs/1908.10084>
- [5] Y. Goldberg and O. Levy, "word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method," *CoRR*, vol. abs/1402.3722, 2014. [Online]. Available: <http://arxiv.org/abs/1402.3722>



# Inteligentné elektromery a ich využitie

Marek GREŠÁK, Dušan MEDVEĎ

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

marek.gresak@student.tuke.sk, dusan.medved@tuke.sk

**Abstrakt** — Článok pojednáva o využití inteligentných elektromerov a ich využití v inteligentnom meracom systéme. V článku sú uvedené inteligentné meracie systémy u nás a vo svete a inteligentné merače používané na Slovensku. V príspevku sú popísané základné charakteristiky inteligentného elektromera. Ďalej je uvedený výpočet spotreby elektriny pri využití rôznych taríf v domácnosti. V závere článku je uvedené zhodnotenie praktickej časti a komplexný pohľad na využitie inteligentných elektromerov v inteligentnom meracom systéme.

**Kľúčové slová** — inteligentný elektromer, inteligentné meracie systémy, spotreba elektrickej energie

## I. ÚVOD

V tomto príspevku sa zameriam na inteligentné meracie systémy, ktoré predstavujú spôsob, ako vylepšiť existujúcu elektrickú sústavu pomocou najnovších výpočtových technológií tak, aby boli schopné čeliť výzvam 21. storočia. To je možné vďaka využívaniu inteligentných meračov inštalovaných v jednotlivých domoch alebo firmách a údajmi, ktoré sa od nich zbierajú v pravidelných krátkych intervaloch. Priblížime si tieto nové inteligentné elektromery a keďže tieto merače poskytujú aj hodnoty o spotrebe elektrickej energie, zahrnieme výpočet spotreby do tohto príspevku a ukážeme si, ako by sa to dalo využiť na úspore financií, či využitia energie pri rozdielnych tarífach elektriny dodávateľov a v rôznych oblastiach, ako sú napríklad domy, firmy či iné spoločnosti.

## II. INTELIGENTNÝ MERACÍ SYSTÉM

Zákazníci sú čoraz viac a viac oprávnení prijímať energetické rozhodnutia, ktoré ovplyvňujú ich každodenný život. Zároveň sa zvyšuje aj dopyt po energii. Účasť spotrebiteľov bude čoskoro zahŕňať rozsiahle používanie elektrických vozidiel (osobných a nákladných automobilov), diaľkové ovládanie domácich spotrebičov na podporu úspory energie a čoraz viac obnoviteľných zdrojov energie s distribuovanou výrobou energie a miestnym riadením skladovania energie. Dostupnosť nových technológií, ako sú distribuované snímače, obojsmerná bezpečná komunikácia, pokročilý softvér na správu údajov a inteligentné a autonómne radiče, otvára nové príležitosti pre transformáciu energetických systémov [1], [2].

### A. Čo je inteligentný merací systém (IMS)

Inteligentný merací systém („IMS“) označuje novú generáciu elektronických systémov, ktoré dokážu merať množstvo elektriny či už spotrebovanej alebo vyrobenej prostredníctvom inteligentných meračov, ktoré môžu zhromažďovať, spracovávať, diaľkovo prenášať a elektronicky poskytovať tieto údaje koncovým zákazníkom. Aby sa zabránilo zneužitiu, údaje sa prenášajú v šifrovanej podobe [3]. Osobné údaje spotrebiteľa sú chránené predpismi EÚ o spracovaní a voľnom pohybe údajov. Inteligentné siete a merače môžu mať vplyv na osobné údaje a súkromie a preto EÚ prijala sériu opatrení na dodržanie pravidiel ochrany údajov. Okrem ochrany údajov a súkromia sa kybernetická bezpečnosť čoraz viac stáva problémom súvisiacim s inteligentnými sieťami a meračmi. Komisia je odhodlaná zmierňovať akékoľvek riziká a zvyšovať odolnosť voči kybernetickej bezpečnosti. Pomocou inteligentných meračov môžu koncoví zákazníci presne a pravidelne získavať nameranú hodnotu svojej spotreby energie a platiť iba za elektrinu, ktorú skutočne používajú. Týmto sa ukončujú nesprávne účty a spätná

fakturácia, ktoré sú v súčasnosti pre spotrebiteľa najväčším záujmom. Inteligentné merače navyše môžu poskytnúť spätnú väzbu o spotrebe energie v reálnom čase a umožniť zákazníkovi, ktorí majú záujem, lepšie spravovať svoje použitie, šetriť energiu a znižovať náklady [1], [4], [5].

### B. IMS na Slovensku

Na území Slovenskej republiky sa inštalácia IMS realizuje na konkrétnom mieste v súlade s povinnosťami prevádzkovateľa distribučnej sústavy podľa zákona č. 251/2012 Z. z. o energetike a vyhlášky č. 358/2013, ktorou sa ustanovuje postup a podmienky v oblasti zavádzania a prevádzky inteligentných meracích systémov v elektroenergetike. Podľa vyššie uvedeného zákona musí prevádzkovateľ distribučnej sústavy nainštalovať inteligentný merací systém v mieste spotreby energie koncového zákazníka pripojeného k distribučnej sústave na nízkej napäťovej úrovni. so spotrebou elektriny najmenej 4 MWh za rok podľa vyhlášky č. 358/2013 Z. z. Starý elektromer (indukčný alebo štandardný statický) viem prevádzkovateľ distribučnej siete bezplatne vymeniť za nový inteligentný. Ak je rozvádzač merača v meracom bode verejne prístupný, zákazník nemusí byť pri inštalácii meradla fyzicky prítomný. Koncoví zákazníci môžu získať informácie o svojich nameraných údajoch priamo z inteligentného merača alebo z webového portálu spoločnosti zo systému IMS. Pomocou týchto informácií môžu vyhodnotiť a aktívne určiť účinnosť spotreby elektriny, čím ovplyvnia ich spotrebu elektriny [4], [6].

### C. IMS vo svete

Infraštruktúra inteligentných sietí a vylepšovanie zariadení sú čoraz populárnejšie. Mnoho miest a krajín zaznamenalo výhody modernizácie. V súčasnosti sa očakáva, že v nasledujúcich desiatich rokoch, 50 krajín investuje 268 miliárd dolárov do infraštruktúry inteligentných sietí, vrátane inteligentných meračov, automatizácie distribúcie a ďalších energetických trhových oblastí. Doteraz sa väčšina investícií do inteligentných sietí sústreďovala pre vyspelé krajiny. Patria sem najmä krajiny v Severnej Amerike, západnej Európe a východnej Ázii.

Tieto trhy tvoria viac ako 75 % aktuálne nainštalovanej základne inteligentných meracích prístrojov, ako aj mnoho priekopníckych automatizačných riešení v oblasti distribúcie energie, analýz, riadenia domácej energie a ďalších programov inteligentnej siete. Mnoho ďalších rozvojových krajín sa však pripravuje na rýchle dobehnutie. Viac ako 50 krajín, medzi ktorými je aj Slovensko, začali skúmať zavedenie inteligentných sietí a v mnohých prípadoch boli ustanovené pokročilé regulačné odbory a rozsiahle pilotné projekty. Infraštruktúra inteligentných sietí prinesie oblastiam, ktoré zavádza, mnoho výhod a bude poháňať tieto krajiny do budúcnosti [7], [8], [9].

## III. INTELIGENTNÉ ELEKTROMERY

Inteligentný elektromer je elektronické zariadenie, ktoré zaznamenáva spotrebu elektrickej energie a mnoho ďalších informácií, pričom komunikuje s dodávateľom elektriny na účely monitorovania a fakturácie. Inteligentné merače zvyčajne zaznamenávajú energiu každú hodinu alebo aj častejšie a komunikujú s dodávateľom najmenej raz za deň.

Inteligentné merače umožňujú obojsmernú komunikáciu medzi meradlom a centrálnym systémom. Takáto pokročilá infraštruktúra sa líši od automatického odpočtu meračov tým, že umožňuje obojsmernú komunikáciu medzi meradlom a dodávateľom.

Komunikácia z meracieho prístroja do siete môže byť bezdrôtová alebo prostredníctvom pevných drôtových pripojení, ako je napríklad elektrické vedenie. Medzi bežné možnosti bezdrôtovej komunikácie patria napr. mobilná komunikácia (ktorá môže byť drahá), Wi-Fi (ľahko dostupné), a rôzne iné bezdrôtové siete. Jedná sa o pokročilú meraciu technológiu zahŕňajúcu umiestňovanie inteligentných meračov na čítanie, spracovanie a spätnú väzbu údajov zákazníkovi. Meria spotrebu energie, na diaľku prepína dodávku zákazníkovi a taktiež na diaľku riadi maximálnu spotrebu elektriny. Inteligentný merací systém využíva pre lepší výkon pokročilú technológiu systému meracej infraštruktúry.

Niektorí ľudia sa obávajú, že inteligentné merače vyžarujú nízkofrekvenčné rádiové vlny, ktoré by mohli byť škodlivé pre ľudí citlivých na elektromagnetické polia. Napriek tomu je to mýtus, ktorý bol rázne popretý. Tieto riziká boli a sú dlhoročne skúmané, pretože čoraz viac domácností bude mať tieto inteligentné merače [10], [11].

## IV. VYUŽITIE DÁT Z INTELIGENTNÉHO ELEKTROMERA NA ŠETRENIE SPOTREBE ELEKTRICKEJ ENERGIE V DOMÁCNOSTI

Jeden zo spôsobov, ako šetriť elektrickú energiu, teda aj peniaze, je využívať spotrebiče v čase, kedy je elektrina najlacnejšia, teda je najmenej odoberaná. Samozrejme, niektoré spotrebiče sa

nedajú využívať iba v špecifickom (predpísanom) čase a preto sme vybrali tie, na ktorých by sa dalo ušetriť. V porovnaní si ukážeme dodávateľa, ktorý má cenu nastavenú po celý deň rovnakú a dodávateľa, ktorý má časové tarify. Pri výpočte sú využívané bežné príkony spotrebičov a priemerný čas ich používania.

A. Dodávateľ elektrickej energie VSD a.s.

CENA ZA DODÁVKU ELEKTRINY OD VSE	Mesačná platba za 1 odberné miesto		Vysoké pásmo (VT) Cena za 1 kWh		Nízke pásmo (NT) Cena za 1 kWh	
	bez DPH EUR	s DPH EUR	bez DPH EUR	s DPH EUR	bez DPH EUR	s DPH EUR
ŠTANDARD* (DD1, DD2)	0,75	0,90	0,0622	0,0746		

Obr. 1 Koncové ceny elektriny pre domácnosti – štandard

Pri výpočte ceny a spotreby energie sa ako vstupné hodnoty využili cena za 1 kWh, bežné príkony spotrebičov a čas použitia, ktorý sa stanovil a pohyboval v rozmedzí 1÷3 hodiny týždenne, podľa vybraného spotrebiča.

V tabuľke nižšie je vypočítaná spotreba elektrickej energie a cena za rok piatich vybraných spotrebičov, s ktorými sa uvažovalo v „šetriacom“ programe.

Tabuľka 1  
 Spotreba energie spotrebičov, vlastné spracovanie spotrebičov, vlastné spracovanie

Nenastaviteľná tarifa			
Spotrebič	Obdobie	Spotreba elektriny (kWh)	Cena (€)
Práčka	1 rok	599,64	44,73
Umýv. riadu	1 rok	938,57	70,02
Vysávač	1 rok	57,36	4,28
Sušička	1 rok	599,64	44,73
Žehlička	1 rok	219	16,34
<b>Cena spolu:</b>			180,1

B. Dodávateľ elektrickej energie Magna Energia a.s.

Tento dodávateľ má cenu nastavenú za 1 kWh podľa časovej tarify ako je zobrazené nižšie v tabuľke.

Tarifa	Časové pásmo	Týždeň [EUR/ kWh]	Víkend [EUR/ kWh]
Noc	22:00 - 06:00	0,03776	0,03096
Ráno/Večer	06:00 - 09:00 19:00 - 22:00	0,07551	0,06192
Predpoludnie	09:00 - 13:00	0,05789	0,04747
Popoludnie	13:00 - 19:00	0,04783	0,03922

Obr. 2 Variabilná tarifkácia

Pri výpočte podľa časovej tarify sme využili všetky časové pásma a taktiež sme ich skombinovali s cenami cez víkend, kedy je elektrická energia lacnejšia, keďže cieľom tohto výpočtu bolo ušetriť čo najviac na spotrebe elektrickej energie daného spotrebiča. Tabuľky na všetky tarify sú zobrazené nižšie.

Tabuľka 2  
Spotreba energie spotrebičov, vlastné spracovanie

<b>Tarifa – ráno/večer/víkend</b>			
<b>Spotrebič</b>	<b>Obdobie</b>	<b>Spotreba elektriny (kWh)</b>	<b>Cena (€)</b>
Práčka	1 rok	599,64	51,07
Umýv. riadu	1 rok	938,57	66,62
Vysávač	1 rok	57,36	3,55
Sušička	1 rok	599,64	42,02
Žehlička	1 rok	219	15,72
<b>Cena spolu :</b>			<b>178,98</b>

Tabuľka 3  
Spotreba energie spotrebičov, vlastné spracovanie

<b>Tarifa – predpoludnie/víkend</b>			
<b>Spotrebič</b>	<b>Obdobie</b>	<b>Spotreba elektriny (kWh)</b>	<b>Cena (€)</b>
Práčka	1 rok	599,64	39,16
Umýv. riadu	1 rok	938,57	50,8
Vysávač	1 rok	57,36	2,72
Sušička	1 rok	599,64	32,22
Žehlička	1 rok	219	11,23
<b>Cena spolu :</b>			<b>136,13</b>

Tabuľka 4  
Spotreba energie spotrebičov, vlastné spracovanie

<b>Tarifa – popoludnie/víkend</b>			
<b>Spotrebič</b>	<b>Obdobie</b>	<b>Spotreba elektriny (kWh)</b>	<b>Cena (€)</b>
Práčka	1 rok	599,64	26,62
Umýv. riadu	1 rok	938,57	42,2
Vysávač	1 rok	57,36	2,25
Sušička	1 rok	599,64	26,62
Žehlička	1 rok	219	9,84
<b>Cena spolu :</b>			<b>107,53</b>

Tabuľka 5  
Spotreba energie spotrebičov, vlastné spracovanie

<b>Tarifa – noc/víkend</b>			
<b>Spotrebič</b>	<b>Obdobie</b>	<b>Spotreba elektriny (kWh)</b>	<b>Cena (€)</b>
Práčka	1 rok	599,64	21,02
Umýv. riadu	1 rok	938,57	33,32
Vysávač	1 rok	57,36	1,78
Sušička	1 rok	599,64	21,02
Žehlička	1 rok	219	7,42
<b>Cena spolu :</b>			<b>84,56</b>

### C. Porovnanie cien dodávateľov elektriny

V tabuľke nižšie si môžeme pozrieť celkové zhrnutie cien podľa určitej tarify a zistiť tak či je

výhodné aj takýmto spôsobom ušetriť financie.

Tabuľka 6  
Celková cena za spotrebu elektrickej energie, vlastné spracovanie

Porovnanie cien		
Dodávateľ	Tarifa	Celková cena
VSD	1kWh	180,1 €
Magna	ráno/večer/víkend	178,98 €
Magna	predpoludnie/víkend	136,13 €
Magna	popoludnie/víkend	107,53 €
Magna	noc/víkend	84,56 €

Ak zhodnotíme všetky 4 tarify, tak je dodávka elektrickej energie s tarifou nastavenou podľa časového pásma pre domácnosti výhodnejšia, pretože je možné ušetriť 70 €, pričom vo výpočte boli zahrnuté len spotrebiče, ktoré by sme mohli najviac využiť v najvýhodnejšej tarife, teda od 13:00 do 19:00.

Využívanie lacnejšej elektriny sa dá využiť aj pri tepelnom čerpadle alebo bojleri, no tieto zariadenia sú už nastavované tak, aby vodu do zásobníka nahrievali v čase, keď je elektrina lacnejšia. Veľa ľudí hľadá výhody inteligentných zariadení a snažia sa šetriť a toto je jedna z možností ako využiť poznatky o spotrebe zariadenia a výhodnej tarife.

## V. ZÁVER

Cieľom tohto príspevku bolo priblížiť si inteligentné elektromery využívané v inteligentnom meracom systéme. Inteligentné merače sú ďalšou generáciou digitálnych elektromerov, ktoré prinesú veľa výhod pre zákazníka, ekonomiku a životné prostredie. Doba napreduje, taktiež aj energetický systém a preto sú a budú inteligentné merače neodmysliteľnou súčasťou v každom meste aj obci. Keďže sa čoraz viac využívajú obnoviteľné zdroje energie, tieto merače budú pomáhať ľuďom nielen pri odbere elektriny, ale aj pri výrobe. Samozrejme sa inteligentné merače nevyužívajú len v elektroenergetike, ale aj pri dodávke plynu a vody.

## POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576, ďalej Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu SR a SAV podporenej grantom VEGA 1/0757/21 a Kultúrnej a edukačnej grantovej agentúre MŠVVaŠ SR pod číslom KEGA 013TUKE-4/2021.

## LITERATÚRA

- [1] Hamid Gharavi, Reza Ghafurian, Smart Grid: *The Electric Energy System of the Future*, [online], [cit.: 7.5.2021] dostupné na <https://ieeexplore.ieee.org/document/5768100>
- [2] Thomas M. Chen, *Smart grids, smart cities need better networks*, [online]; [cit.: 8.5.2021] dostupné na <https://ieeexplore.ieee.org/abstract/document/5430136>
- [3] Super savers: *Meters to manage the future*, [online], [cit.: 7.5.2021] dostupné na <https://www.nature.com/articles/445586a>
- [4] Východoslovenská distribučná, a.s.: *Inteligentné meracie systémy*, [online], [cit.: 9.5.2021] dostupné na <https://www.vsd.sk/edso/domov/technicke-info/meranie-distribucie/ims>
- [5] European commission: *Smart grids and meters*, [online], [cit.: 7.5.2021] dostupné na [https://ec.europa.eu/energy/topics/markets-and-consumers/smart-grids-and-meters\\_en?redir=1](https://ec.europa.eu/energy/topics/markets-and-consumers/smart-grids-and-meters_en?redir=1)
- [6] Elektrometer – *Čo je to IMS?*, [online], [cit.: 9.5.2021] dostupné na <https://www.elektromer.online/ims>
- [7] Matt Wakefield, Bartosz Wojszczyk: *Smart Grids*, 2nd edition, chapter Global Smart Grid Initiatives, 2013
- [8] IOT Analytics - Smart Meter Market 2019: *Global penetration reached 14% – North America, Europe ahead*, [online], [cit.: 7.5.2021] dostupné na <https://iot-analytics.com/smart-meter-market-2019-global-penetration-reached-14-percent/>
- [9] MI Smart Grids, *Smart Grids Innovation Challenge Country Report 2019 - Strategies, Trends and Activities on Jointly Identified Research Topics (START)*, [online], [cit.: 9.5.2021] dostupné na [https://smartgrids.no/wp-content/uploads/sites/4/2019/08/2019\\_MI\\_IC1\\_Country\\_Report.pdf](https://smartgrids.no/wp-content/uploads/sites/4/2019/08/2019_MI_IC1_Country_Report.pdf)
- [10] TZB-info - *Elektronické statické elektroměry – seznamte se*, [online], [cit.: 7.5.2021] dostupné na <https://elektro.tzb-info.cz/13933-elektronicke-staticke-elektromery-seznamte-se>
- [11] The Spruce: *How an Electric Meter Reads Power Usage*, [online], [cit.: 9.5.2021] dostupné na <https://www.thespruce.com/how-electric-meters-read-power-1152754>
- [12] Grešák, M.: *Inteligentné elektromery a ich využitie*. Bakalárska práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2021.

# Dielectric spectroscopy of transformer paper in the time domain

<sup>1</sup>Peter Havran, <sup>2</sup>Roman Cimbala, <sup>3</sup>Luboš Šárpataky

<sup>1</sup> Department of Electric Power Engineering, 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>3</sup> Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical university of Košice, Slovak republic

<sup>1</sup>peter.havran@tuke.sk, <sup>2</sup>roman.cimbala@tuke.sk, <sup>3</sup>lubos.sarpataky@tuke.sk

**Abstract** — This publication addresses the description of electrophysical changes in transformer paper that is subjected to accelerated thermal aging. The introduction includes an explanation of the importance of transformer paper in the power system and a description of a non-destructive method of dielectric spectroscopy in the time domain, which will be used to analyze changes in the thermal degradation process. The experiment contains a description of the measuring workplace, the implementation procedure, and the analysis of the measured data with an explanation of the findings.

**Keywords** — charging current, dielectric spectroscopy, paper, thermal aging

## I. INTRODUCTION

The electricity industry is a major pillar in the development of civilized human society. The operability of its elements, such as transformers, depends on the reliability of the operation of energy equipment. Power transformers are unpredictable due to the aging of their insulation system, which forms a paper-impregnated composition [3].

The operation of the transformer results in thermal aging of the insulating paper, which reduces the life of the transformer. As a result of thermal aging, substances, water, and organic acids are formed, which accelerate the aging of the transformer paper. Therefore, it is necessary to focus on improving the insulating properties of the transformer and improve its life with new materials [1]. The life of the transformer depends on the life of the insulation paper. According to world statistics, less than 5% of transformers have a service life of more than 30 years and about 1% for 40 years. Therefore, it is necessary to examine the paper insulation in transformers. Its use is preferred due to its low cost and ease of manufacture, but in terms of thermal stability, it is not very suitable because at certain temperatures the molecular chain is degraded and the mechanical strength is reduced. Much research has been devoted to the analysis of transformer paper aging, such as furfural formation, mechanical strength, and the analysis of the degree of polymerization, which is important for determining the insulation state of transformer paper [2].

Also, an important method that analyzes the state of insulating paper at the molecular level is time-domain dielectric spectroscopy. This method examines polarization processes in the time domain when applying a DC electric field. By applying a direct electric field, the dielectric is charged, resulting in a characteristic of the charging current over time. With the subsequent loss of the electric field, the dielectric discharges by means of a discharge current over time. The charging (polarizing) and discharging (depolarizing) currents have an exponential curve, which is the exponential sum of the individual polarization processes occurring. The difference between the polarization and depolarization currents lies in the existence of a conductivity process during charging, i.e. the polarization current also contains the contribution of the conductivity current. By means of this method, a transformer paper will be analyzed in this publication, which will be investigated in the process of accelerated thermal aging, indicating changes in the electrophysical structure of the material.

## II. EXPERIMENT

The accelerated thermal aging experiment was performed on circular samples of transformer paper with a diameter of 74 mm and a thickness of 0.42 mm. Fifteen such samples were prepared because 3 samples were measured at each stage of thermal aging. The experiment contains five degrees of thermal aging at 160 °C with an exponential increase according to Arrhenius's theory of thermal aging. A simulation of the degree of thermal aging is shown in Fig. 1.

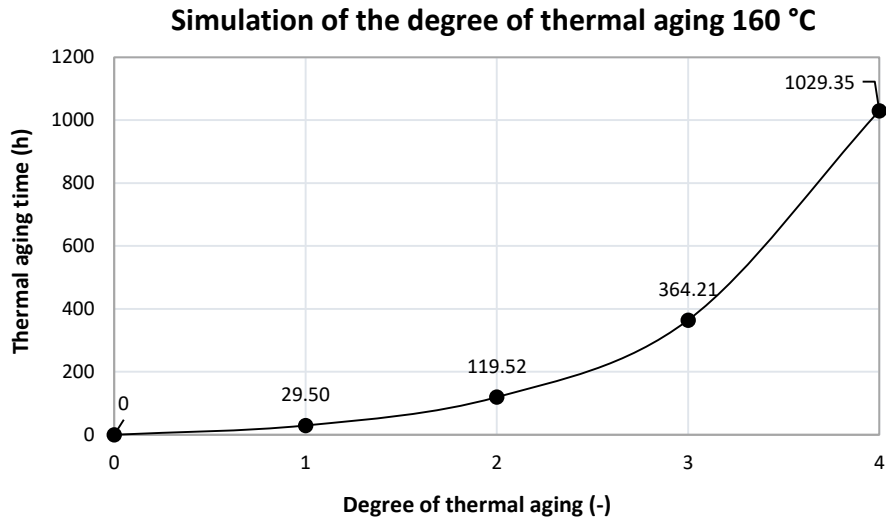


Fig. 1 Simulation of the degree of thermal aging at 160 °C

Thermal aging was performed in a UF55<sup>plus</sup> universal oven. After preparing 15 paper samples, 12 paper samples were placed in a Petri dish, which was then placed in a heating oven at a set temperature of 160 °C. The first three samples were not placed in the oven because they represented the measurement at zero degree of thermal aging (0 h). After measuring the zero degree of aging, the samples were stored in postal envelopes and stored securely. After a first heat aging stage of 29.5 hours, three samples were taken from the Petri dish for measurement at 29.5 hours. The other nine paper samples in the Petri dish were put back in the oven. After measuring the first stage of aging, the samples were also placed in postal envelopes and stored securely. The whole experiment was carried out with the same procedure until the thermal aging time of 1029.35 h.

Time-domain dielectric spectroscopy was performed using a Keithley 617 measuring instrument, which was connected via GPIB bus to a desktop computer with built-in Agilent VEE Pro software. The software recorded the values of the charging current in a time zone of 1000 s at an applied DC voltage of 100 V. The paper samples were measured in a Tettex Ag Zürich electrode system type 2904 (Guard Ring Capacitor), which was connected by wires to a measuring instrument. The wiring diagram with a description of the individual devices is shown in Fig. 2.

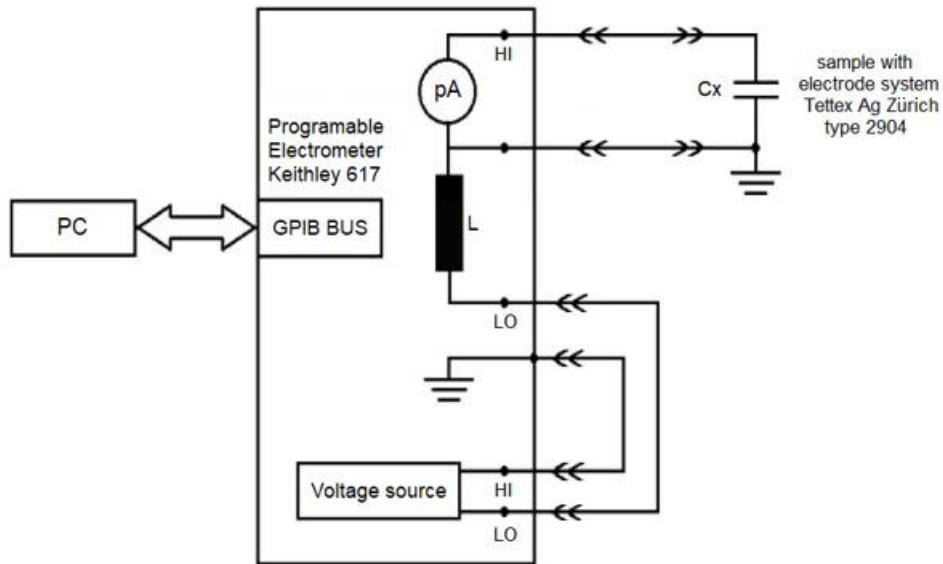


Fig. 2 Wiring diagram of experimental measurement

### III. ANALYSIS OF ACCELERATED THERMAL AGING

In this part, the values of charging currents in the time spectrum at accelerated thermal aging of 1029.35 h at a temperature of 160 °C of three samples of transformer paper will be analyzed. Fig. 3 shows the measured data of the first paper sample. The thermal aging of 29.5 h did not bring about a significant change in the charging current courses at the time of the first sample. During the thermal aging time of 119.52 h, there was a significant increase in the polarization and conductivity current, which can be expected with increasing thermal aging. Increasing the thermal aging to 364.21 h reduced the polarization current to below the thermal aging values of 0 h, which may be due to the hardening of the insulating paper, because by increasing the thermal aging time the paper improved its insulating properties. As for the conductivity current in the small window in Fig. 3, thermal aging of 364.21 h caused an increase in the conductivity of the paper compared to 0 h and 29.5 h. During the thermal aging time of 1039.35 h, the polarization current decreased again and there was no significant change in the conductivity current.

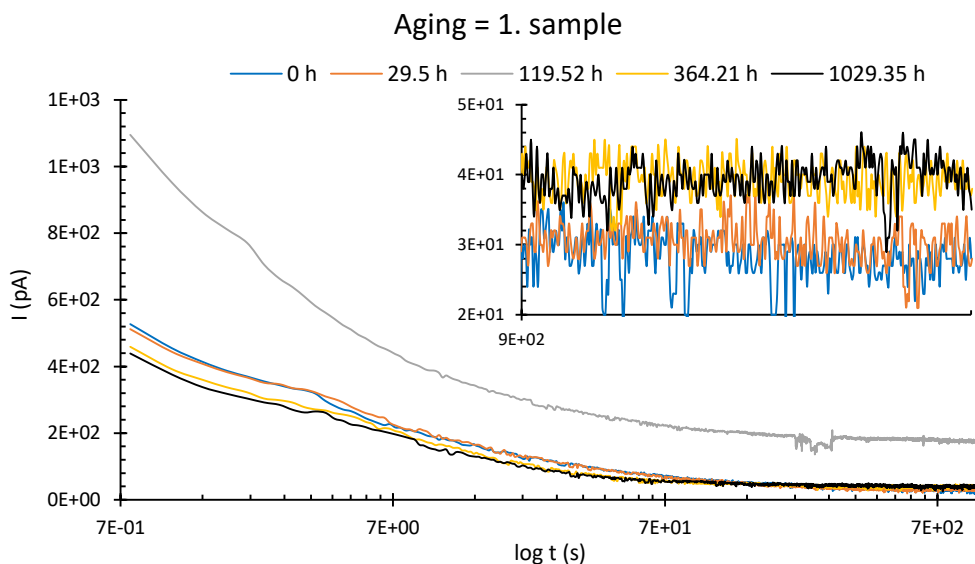


Fig. 3 Time dependence of the charging current during accelerated thermal aging of the 1st paper sample

Fig. 4 shows the time dependence of the charging current of the second sample. Thermal aging of 29.5 h resulted in a slight increase in polarization current with a slight decrease in conductivity current compared to 0 h. As with the first sample, there was a significant increase in charging current during the thermal aging time of 119.52 h. At a thermal aging time of 364.21 h, the polarization current was shifted below the characteristic of 0 h and the conductivity current was again higher similarly to the first sample. The thermal stability of the



insulating paper remained unchanged even at 1029.35 h, but the conductivity current was reduced, which has the effect of slightly improving the insulating properties in this respect.

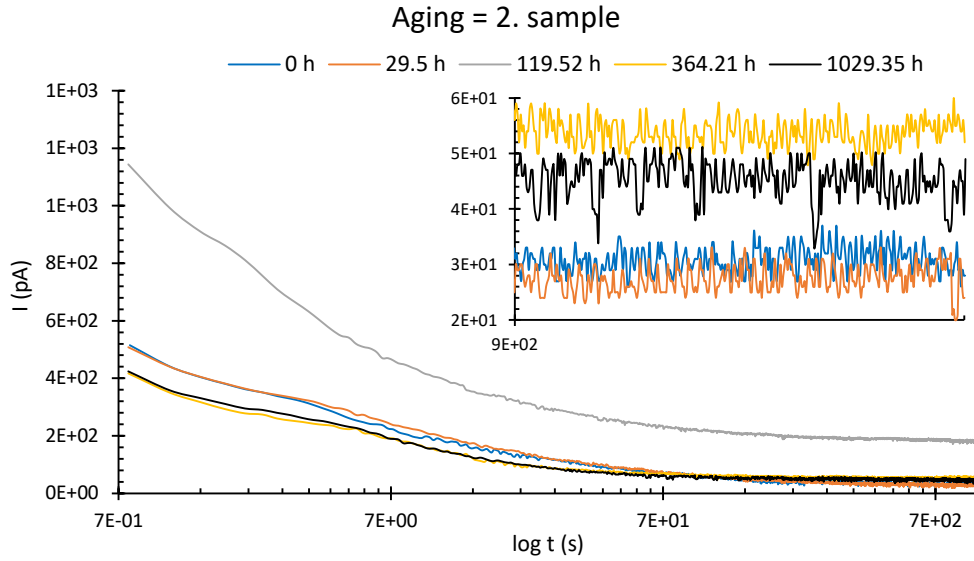


Fig. 4 Time dependence of the charging current during accelerated thermal aging of the 2nd paper sample

The charging currents in the time spectrum of the third sample in Fig. 5 have certain conclusive conclusions that can be expressed from this experiment. The thermal aging time of 29.5 h again did not contribute to significant changes in the insulating state of the transformer paper. As with the first and second samples, the thermal aging time of 119.52 h represents a significant increase in the charging currents of the samples. With an exponential increase in thermal degradation, the charging currents decrease at 364.21 h, i.e. 119.52 h represents the cut-off value for the start of the hardening process of the transformer paper at 160 °C. Thermal aging of 1029.35 h caused an increase in polarization current in the third sample compared to 364.21 h, but on the other hand, it had a lower conductivity current similar to the second sample.

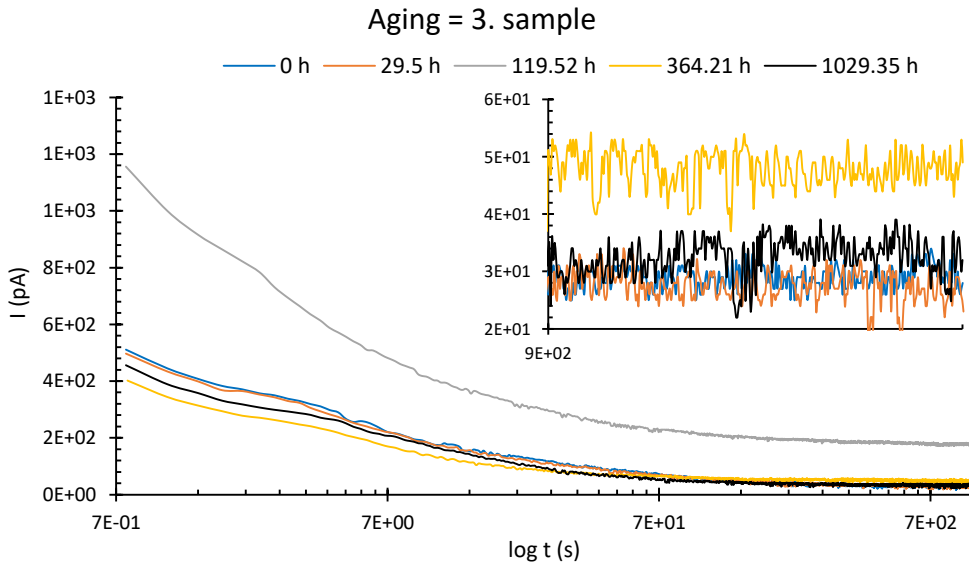


Fig. 5 Time dependence of the charging current during accelerated thermal aging of the 3rd paper sample

#### IV. CONCLUSION

This publication included an accelerated thermal aging experiment of transformer paper, which was analyzed by a non-destructive time domain dielectric spectroscopy method. The following conclusions can be drawn from the experiment. A thermal aging time of 119.52 h at 160 °C causes the hardening process of the transformer paper to begin. This is confirmed by the reduction of charging currents with an exponentially increasing thermal aging time of 364.21 and 1029.35 h. The polarization and conductivity currents did not have the same ratios, because the time of 364.21 h reached lower values of the polarization current and, conversely, higher

values of the conductivity current than the thermal aging time of 1029.35 h. This experiment showed that a thermal aging time of 1029.35 h caused a lower polarization current and a higher conductivity current than thermally undamaged samples (0 h). Therefore, it can be argued that at this thermal aging time of 160 °C, the insulating properties of the transformer paper do not change significantly. To describe the more significant changes taking place in these samples, it is necessary to examine the paper insulation with these conditions further, by exponentially increasing the thermal aging time above our investigated value of 1029.35 h.

#### ACKNOWLEDGMENT

The work was supported by the Ministry of Education, Science, Research and Sport within the project VEGA 2/0011/20 and 1/0154/21 and the Slovak Agency for Research and Development based on contract no. APVV-15-0438, APVV-17-0372, and APVV-18-0160.

#### REFERENCES

- [1] C. Tang, S. Zhang, X. Li, B. Xiong, J. Xie, „Experimental analyses and molecular simulation of the thermal aging of transformer insulation paper“, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 22, No. 6, 2015. 3608-3616 pp. ISSN 1558-4135.
- [2] L. Ruijin, T. Chao, Y. Lijun, C. Huanchao, „Thermal Aging Studies on Cellulose Insulation Paper of Power Transformer Using AFM“, 2006 IEEE 8<sup>th</sup> International Conference on Properties & applications of Dielectric Materials, 2007. 722-725 pp. ISSN 2160-9241.
- [3] N. Zhuravleva, A. Reznik, A. Tukacheva, D. Kiesewetter, E. Smirnova, „The study of thermal aging components paper-impregnated insulation of power transformers“, 2016 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EConRusNW), 2016. 747-751 pp. ISBN 978-1-5090-0445-4.

# Porovnanie transformačných operátorov horolezeckého algoritmu v rámci domény TSP

<sup>1</sup>Jozef FRIVALDSKÝ, <sup>2</sup>Marián MACH

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

<sup>1</sup>jozef.frivaldsky@student.tuke.sk, <sup>2</sup>marian.mach@tuke.sk

**Abstrakt** — Článok sa zameriava na problematiku riešenia úloh kombinatorickej optimalizácie s dôrazom na riešenie problému obchodného cestujúceho (TSP) využitím lokálnych prehľadávacích algoritmov. Pozornosť je zameraná na porovnanie transformačných operátorov, používaných v rámci lokálnych prehľadávacích algoritmov pre transformáciu cesty v grafe na inú cestu. Článok sa zaoberá štyrmi základnými operátormi (REVERSE, SWAP, INSERT, GSR) a jedným zovšeobecňujúcim operátorom (3-OPT). Kľúčovými faktormi vzájomného porovnania sú časové nároky použitia týchto operátorov a kvalita dosahovaných riešení. V rámci experimentovania bolo realizovaných niekoľko experimentov s týmito operátormi v rámci domény TSP na vybranom testovacom probléme.

**Kľúčové slová** — horolezecký algoritmus, lokálne prehľadávanie, transformačný operátor, TSP

## I. ÚVOD

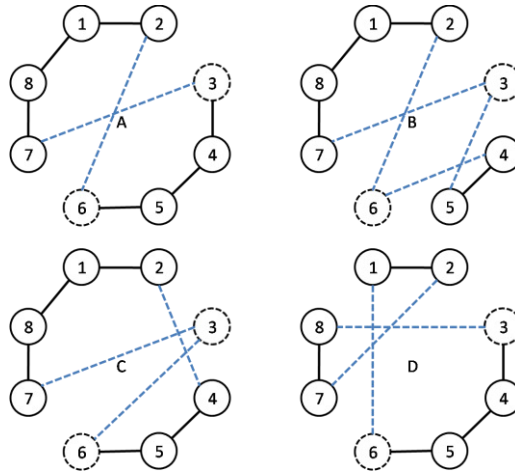
Problém obchodného cestujúceho TSP (Traveling Salesman Problem) je známym problémom kombinatorickej optimalizácie. Formálne je definovaný ako grafová úloha – je zadaná množina uzlov a vážených hrán medzi uzlami, pre ktoré je potrebné nájsť najkratšiu uzavretú cestu - Hamiltonovskú kružnicu (kruhovú cestu prechádzajúcu každým uzlom práve raz). Problém patrí do NP-hard zložitejšej triedy [1]. Potreba riešiť úlohu TSP je pomerne častou v každodennej praxi, pretože sa jedná o prototypový problém modelujúci veľa praktických úloh.

Jednou z populárnych skupín metód riešenia TSP sú prehľadávacie algoritmy, vychádzajúce z existencie priestoru prehľadávania, tvoreného veľkým počtom návrhov na hľadané riešenie (v prípade TSP Hamiltonovských kružníc rôznej dĺžky). Tento priestor je prehľadávaný so snahou nájsť taký návrh, ktorý je najpriateľnejší (najkratšia z Hamiltonovských kružníc). Podtriedou takýchto algoritmov sú algoritmy lokálneho prehľadávania [2], ktoré sa vzdávajú garancie nájdenia optimálneho riešenia výmenou za prijateľné časové nároky, umožňujúce ich praktické použitie. Pri prehľadávaní generujú jednotlivé návrhy riešení v určitom poradí. Ťažiskom prehľadávania je výber nasledujúceho návrhu, rozhodnutie ktorá cesta v grafe má byť preskúmaná najskôr. Tento výber sa nedeje z celého priestoru prehľadávania ale iba z jeho časti – aby nový návrh bol podobný aktuálne skúmanému návrhu. Preto sa vyberá iba z lokálneho okolia aktuálneho návrhu. Prakticky toto okolie je dané ako množina všetkých tých návrhov riešenia, na ktoré je možné aktuálny návrh transformovať pomocou nejakého transformačného operátora. Transformačné operátory sú doménovo špecifické a preto sa budeme venovať takým transformačným operátorom, ktoré sú schopné pretransformovať jednu cestu v grafe na inú cestu pri dodržaní podmienky, že pôvodná aj nová cesta sú Hamiltonovské kružnice a obe cesty sú navzájom podobné.

## II. TRANSFORMAČNÉ OPERÁTORY

V doméne TSP sa používajú transformačné operátory na transformáciu jednej cesty na cestu inú, pričom výsledná cesta si zachováva veľký stupeň podobnosti s pôvodnou cestou. Podstatou transformácie je vykonaním malej perturbácie pôvodnej cesty získať zmenenú cestu. Na túto transformáciu je možné nahliadať ako na zmenu množiny hrán (pretože uzly sa nemenia, mení sa iba ich poradie definované použitými hranami) – každá transformácia má za následok

vypustenie určitého počtu hrán a pridanie rovnakého počtu nových hrán. Typicky sa jedná o malé počty – dve až štyri hrany. Princiálne je možné vykonať aj masívnejšie zmeny použitím väčšieho počtu menených hrán, tým sa však znižuje podobnosť pôvodnej a novej cesty. Zameriame sa na štyri transformačné operátory – operátory REVERSE (označovaný aj ako 2-opt operátor) [3], SWAP, INSERT [4] a GSR (Greedy Subtour Reconnection) [5]. Ich činnosť je ilustrovaná na obr. 1, kde došlo k perturbácii cesty zlozenej z ôsmich hrán 1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8 a 8-1.



Obr. 1 Ilustrácia pôsobenia transformačných operátorov REVERSE (A), SWAP (B), INSERT (C) a GSR (D).

Všetky štyri operátory vyžadujú náhodný výber dvoch uzlov (na obrázku bol najprv vybraný uzol 3 a potom uzol 6). Ďalšia činnosť operátorov sa už líši:

- REVERSE – v úseku definovanom vybranými dvomi uzlami (vrátane týchto dvoch uzlov) sa poradie uzlov obráti (obr. 1 - A). Dôsledkom je, že v ceste sa dve hrany zamenia za dve nové hrany.
- SWAP – vybrané uzly si navzájom vymenia svoju polohu (obr. 1 - B). Dôsledkom je, že v ceste sa štyri pôvodné hrany zamenia za štyri nové hrany.
- INSERT – prvý vybraný uzol sa zo svojho umiestnenia v ceste presunie tesne za druhý vybraný uzol (obr. 1 - C). Dôsledkom je, že v ceste sa tri hrany zamenia za tri nové hrany.
- GSR – úsek definovaný vybranými dvomi uzlami (vrátane tých uzlov) sa vyberie zo svojej pozície v ceste a nahradí niektorú hranu v zostávajúcej časti cesty (obr. 1 - D). Dôsledkom je náhrada troch hrán.

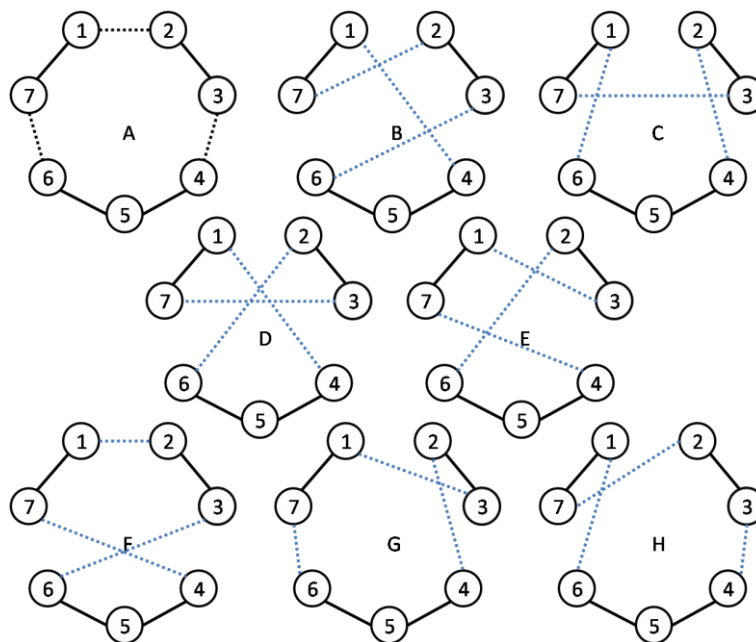
Masívnosť zmien použitím jednotlivých operátorov je rôzna – najmenšie zmeny má za následok operátor REVERSE zatiaľ čo cestu najmasívnejším spôsobom mení operátor SWAP.

Prípado dvojitého výberu toho istého uzla má zmysel iba pre operátor GSR. Cesta identická k zdrojovej ceste je zbytočne vytváraná operátormi REVERSE (ak zamieňané hrany incidujú s rovnakým uzlom, obrátenie poradia tohto uzla hrany nezmení) a SWAP (uzol nahrádza sám seba). Avšak je neprípustným prípadom pre operátor INSERT (uzol sa v ceste nemôže vyskytovať dvakrát).

Ak vybrané dva uzly navzájom v ceste susedia, tak potom operátory SWAP a INSERT degradujú (namiesto výmeny troch alebo štyroch hrán dochádza k výmene iba dvoch hrán) a produkujú rovnakú výslednú cestu ako operátor REVERSE.

Zovšeobecnením operátora REVERSE je  $k$ -opt operátor, ktorý nahrádza  $k$  pôvodných hrán za nové hrany. Čím je tento počet väčší, tým masívnejšie zmeny dokáže operátor realizovať. Operátor 3-OPT pracuje s trojicou hrán. Keďže tri dvojice uzlov je možné prepojiť ôsmimi spôsobmi, tak operátor principiálne z jednej cesty je schopný vygenerovať sedem ďalších ciest. Je to znázornené na obr. 2. Na tomto obrázku A označuje zdrojovú cestu. Je zrejme, že prípady B, C, D a E reprezentujú masívnejšiu náhradu troch hrán, zatiaľ čo prípady F, G a H reprezentujú menej rozsiahle zmeny, pretože degradujú iba na náhradu dvoch hrán (tretia hrana bola nahradená sama sebou). Pri realizácii operátora je potrebné rozhodnúť, či sa generuje

- iba jedna nová cesta (teda sa vyberie iba jedna z možných sedem alternatív) alebo sa generuje viac ciest
- iba z tých alternatív, ktoré menia tri hrany (štyri alternatívy) alebo z alternatív meniacich iba dve hrany (tri alternatívy) alebo zo všetkých (sedem) alternatív.



Obr. 2 Ilustrácia pôsobenia transformačného operátora 3-OPT pri zmene troch hrán.

Rozhodli sme sa použiť štyri varianty operátora. Tie, ktoré generujú iba jednu alternatívu sú označené „RND“ (pretože tá alternatíva sa vyberá náhodne spomedzi uvažovaných alternatív). Tie varianty, ktoré generujú všetky uvažované alternatívy, sú označené „BEST“ (pretože generované alternatívy sú navzájom porovnávané a iba najlepšia je ponechaná zatiaľ čo ostatné sú zahodené). Varianty, ktoré uvažujú všetkých sedem možných alternatív, sú označené „full“ zatiaľ čo varianty uvažujúce iba štyri nedegradované alternatívy sú bez označenia (tri degradované alternatívy už generoval operátor REVERSE). Vznikli tak štyri varianty operátora

- 3-OPT RND – náhodne generuje jednu zo štyroch alternatív
- 3-OPT RND-full – náhodne generuje jednu zo siedmich alternatív
- 3-OPT BEST – generuje štyri alternatívy a vyberá z nich najlepšiu
- 3-OPT BEST-full – generuje sedem alternatív a vyberá z nich najlepšiu

ktoré boli použité v realizovaných experimentoch.

### III. VYBRANÉ EXPERIMENTY S TRANSFORMAČNÝMI OPERÁTORMI

#### A. Testovací problém

Pre experimentovanie s rôznymi transformačnými operátormi bol pre samotné testovanie použitý problém „Berlin52“ z TSPLIB (knížnica inštancií rôznych typov TSP) [6]. Problém je definovaný nad 52 uzlami, pričom je symetrický (nezáleží na smere prechodu medzi uzlami) a Euklidovský (spôsob určovania vzdialenosti). Hľadaným riešením je cyklická cesta cez všetkých  $n$  uzlov, pričom optimalizačnou funkciou je dĺžka cesty. Priestor prehľadávania pozostáva z

$$\frac{n!}{2n} = \frac{52!}{2 \cdot 52} = 7,755593766 \cdot 10^{65} \quad (1)$$

ciest, pričom každá z nich bude reprezentovať jednu Hamiltonovskú kružnicu. Aktuálne sa za najkratšiu cestu považuje cesta s dĺžkou 7542.

#### B. Metodika testovania

Pre testovanie bol použitý horolezecký algoritmus (známy aj pod menom iteračne zlepšovanie), ktorý je prototypovým príkladom lokálneho prehľadacieho algoritmu [7]. Algoritmus štartuje prehľadávanie z náhodne vygenerovanej cesty a následne pomocou opakovaného aplikovania transformačného operátora podľa výberovej stratégie prvého náhodného zlepšenia [8] prechádza ďalšími cestami takým spôsobom, že hodnota optimalizačnej funkcie pre každú novo vybranú cestu je lepšia ako hodnota predchádzajúcej (teda dĺžky ciest sa postupne znižujú). Ak už nie je schopný nájsť cestu s lepšou hodnotou optimalizačnej funkcie, tak svoju činnosť ukončí – hľadanie končí uviaznutím v lokálnom extréme, ktorý môže ale nemusí byť extrémom globálnym. Kvalita tohto lokálneho extrému sa

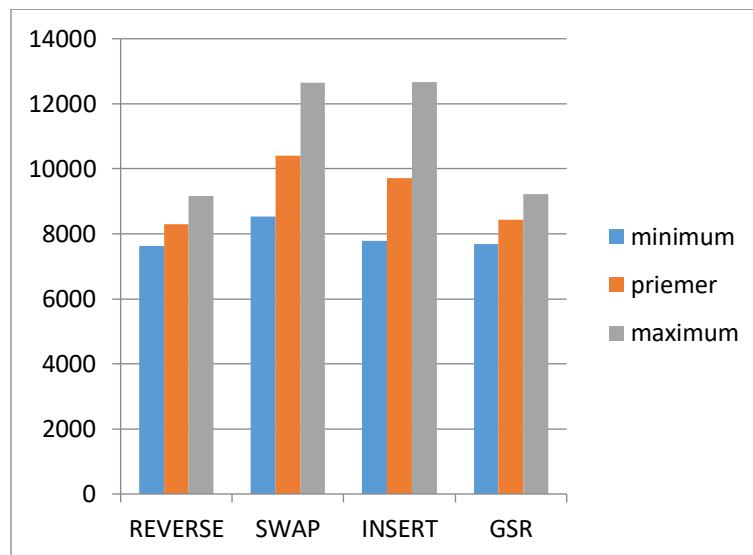
nedá vopred určiť ani garantovať.

Pri testovaní bola využitá softvérová implementácia z [9]. Z dôvodu stochastickej povahy horolezeckého algoritmu bolo pre každý transformačný operátor použité opakované hľadanie s použitím 1000 opakovaní.

Pre meranie kvality nájdených riešení (ciest) bola priamo použitá dĺžka týchto ciest. Pre meranie časovej náročnosti nebol použitý priamo čas, pretože ten by silne závisel ako na charakteristikách hardvéru použitého pre testovanie tak aj na kvalite softvérovej implementácie horolezeckého algoritmu a použitých operátorov. Namiesto toho boli použité dva zástupné parametre – počet iterácií algoritmu potrebných pre to aby sa algoritmus z iniciálneho stavu dopracoval k uviaznutiu v lokálnom extréme a celkový počet generovaných ciest. Vychádzali sme z toho, že zložitosť použitých operátorov (okrem 3-OPT) je porovnateľná (generujú rovnako veľké okolia, keďže vychádzajú z výberu dvoch uzlov v aktuálnej ceste).

### C. Porovnanie transformačných operátorov

Výsledky experimentov s jednotlivými transformačnými operátormi z hľadiska dĺžky nájdených ciest sú uvedené na obr. 3. Porovnávané sú operátory REVERSE (najlepšia nájdená cesta 7619), SWAP (najlepšia cesta 8526), INSERT (najlepšia cesta 7793) a GSR (najlepšia cesta 7689).



Obr. 3 Dĺžky ciest generovaných operátormi

Z obrázka a údajov o najlepšej nájdenej ceste je zrejmé, že operátor SWAP zaostáva za ostatnými operátormi. Operátory REVERSE a GSR sú porovnateľné ako z hľadiska kvality najlepšieho nájdeného riešenia tak aj z hľadiska robustnosti (konzistentnosti nachádzania dobrých riešení). Operátor INSERT za nimi o niečo zaostáva z hľadiska najlepšieho riešenia. Robustnosť je problémom pre oba operátory SWAP a INSERT. Z hľadiska kvality je teda možné doporučiť operátory REVERSE a GSR.

Výsledky časových charakteristík sú uvedené v tab. 1. Tabuľka obsahuje dvojice hodnôt, kde prvá hodnota udáva počet iterácií algoritmu a druhá celkový počet generovaných ciest. Základné operátory sú uvedené v ľavej polovici tabuľky.

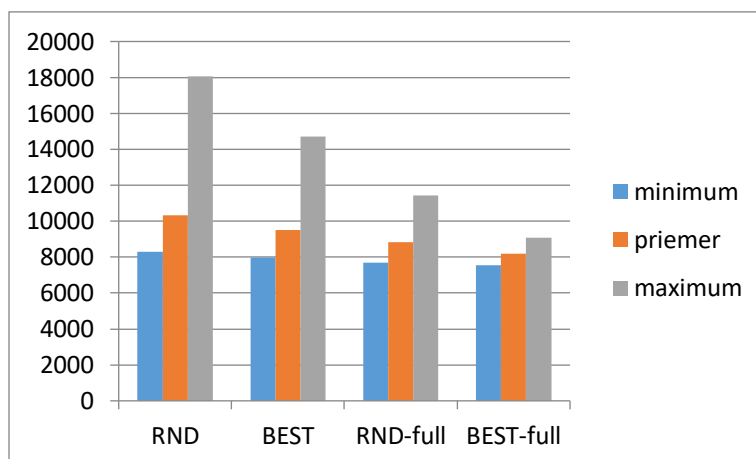
Vyššie favorizované operátory REVERSE a GSR aj z hľadiska počtu generovaných ciest (a tým aj z hľadiska časovej zložitosti) sa javia lepšou voľbou oproti ostatným dvom operátorom SWAP a INSERT. Tie síce dosahujú porovnateľné minimálne hodnoty, avšak v priemerných hodnotách už zaostávajú a maximálne hodnoty, skoro dvojnásobné oproti operátoru REVERSE, ich odsúvajú z pozície vhodných kandidátov na použité operátory pre prípad, že nejaký operátor má byť použitý samostatne.

Na pravej strane tab. 1 sú uvedené hodnoty pre štyri varianty operátora 3-OPT – varianty RND (najlepšia nájdená cesta 8304), BEST (najlepšia cesta 7975), RND-full (najlepšia cesta 7688) a BEST-full (najlepšia cesta 7544). Ich porovnanie z hľadiska dĺžok generovaných ciest je na obr. 4.

Tabuľka 1

Časová náročnosť hľadania riešenia rôznymi transformačnými operátormi. Prvá hodnota v dvojiciach reprezentuje počet realizovaných iterácií algoritmu, druhá hodnota je počet generovaných ciest.

Operátor	Náročnosť	Operátor	Náročnosť		
REVERSE	Min	97 / 3970	Min	50 / 1722	
	Priemer	142 / 6991	3-OPT	Priemer	122 / 7234
	Max	189 / 11894	RND	Max	190 / 20607
SWAP	Min	86 / 4338	Min	51 / 2675	
	Priemer	131 / 9397	3-OPT	Priemer	124 / 19760
	Max	195 / 22097	BEST	Max	218 / 63919
INSERT	Min	86 / 3717	Min	100 / 3561	
	Priemer	128 / 8242	3-OPT	Priemer	152 / 10352
	Max	183 / 21358	RND-full	Max	214 / 22910
GSR	Min	64 / 2333	Min	109 / 13395	
	Priemer	91 / 5960	3-OPT	Priemer	149 / 36231
	Max	128 / 16937	BEST-full	Max	204 / 83913



Obr. 4 Dĺžky ciest generovaných variantmi operátora 3-OPT

Horšie výsledky variantov RND a BEST sú spôsobené tým, že tieto varianty robia masívnejšie zmeny v generovaných cestách a preto zlyhávajú v konečnej fáze jemného doladovania riešenia. Navyše ich počty generovaných ciest sa pohybujú v značnom rozsahu. Variant RND-full dosahuje veľmi dobré výsledky z hľadiska kvality najlepšej cesty, ale spája to so strednou robustnosťou kvality riešení a vyššou náročnosťou reprezentovanou priemerným počtom generovaných ciest. Najzaujímavejším variantom je BEST-full, ktorý bol schopný nájsť najkratšiu cestu avšak platí za to enormným nárastom počtu generovaných ciest – čo je spôsobené tým, že generuje vždy viac (sedem) ciest z ktorých si iba jednu ponechá a ostatné zahodí. Pri snahe nájsť čo najkvalitnejšie riešenie pri dostatku času na jeho hľadanie je operátor 3-OPT BEST-full favoritom.

V snahe zvýšiť kvalitu nájdených riešení je možné namiesto iba jedného operátora použiť operátorov viacero s tým, že budú nejakým spôsobom kombinované. Použili sme pravdepodobnostnú kombináciu – pri generovaní novej cesty sa náhodne vybral jeden operátor z množiny dostupných alternatív, pričom pri tomto výbere nebol preferovaný žiadny z operátorov. Ako alternatívy boli použité iba operátory REVERSE, SWAP, INSERT a GSR (3-OPT sme nekombinovali, pretože variant BEST-full našiel riešenie veľmi blízko globálneho extrému testovacieho problému a teda nemožno očakávať ešte ďalšie zlepšenie). Výsledky kombinovania operátorov do dvojíc sú v tab. 2.

Tabuľka 2

Najlepšie riešenia nájdené kombináciami dvojíc rôznych transformačných operátorov. Druhá hodnota v dvojiciach reprezentuje dĺžku najkratšej nájdennej cesty, prvá hodnota koľkokrát bola táto cesta nájdená z 1000 behov algoritmu.

Operátor	REVERSE	SWAP	INSERT	GSR
REVERSE		2 / 7672	14 / 7544	16 / 7544
SWAP			1 / 7767	1 / 7585
INSERT				3 / 7544

Z tabuľky je zrejmé, že operátor SWAP sa hodí do roly jednej z alternatív na kombinovanie menej než ostatné tri operátory. Dvojice vyberané z operátorov REVERSE, INSERT a GSR sa dostali rovnako blízko globálnemu extrému ako 3-OPT BEST-full, avšak každá z troch možných kombinovaných dvojíc generuje niekoľkokrát menej ciest než spomínaný operátor 3-OPT BEST-full a teda pri rovnakej kvalite riešenia je potrebný kratší čas na jeho hľadanie. Toto robí favorita z používania kombinácie operátorov pred použitím iba jedného operátora.

Tabuľka 3

Najlepšie riešenia nájdené kombináciami trojíc a štvoric rôznych operátorov. Druhá hodnota v dvojiciach reprezentuje dĺžku najkratšej nájdennej cesty, prvá hodnota koľkokrát bola táto cesta nájdená z 1000 behov algoritmu.

REVERSE	SWAP	INSERT	GSR	Dĺžka cesty
X	X	X		9 / 7544
X	X		X	32 / 7544
X		X	X	40 / 7544
	X	X	X	7 / 7544
X	X	X	X	25 / 7544

Tab. 3 zobrazuje výsledky pre prípad kombinovania viacerých než dvoch operátorov – opäť na pravdepodobnostnom základe bez preferencií. Ako vidno, kvalita riešenia sa už nezlepšila, avšak zlepšila sa frekvencia nachádzania toho najlepšieho riešenia – pre kombináciu operátorov REVERSE, INSERT a GSR je najlepšia cesta nachádzaná v 4% vykonaných pokusov.

#### IV. ZÁVER

Vykonané experimenty potvrdili, že medzi transformačnými operátormi existujú rozdiely a je preto pri plánovaní ich použitia potrebné zvážiť, čo je kontextom zamýšľaného hľadania. Ak je dôraz kladený na dobu hľadania s akceptovateľnosťou aj o niečo horšieho riešenia, tak je možné odporučiť samostatné použitie operátora REVERSE alebo GSR. Ak je dôraz na čo najväčšiu kvalitu riešenia bez ohľadu na časové hľadisko, potom sa vhodným odporúčaním javí operátor 3-OPT vo variante BEST-full, pretože možnosť vykonávať aj masívnejšie transformačné zmeny poskytuje šancu na únik z lokálneho extrému v prípade, keď menej masívne zmeny na únik nestačia. V prípade nutnosti kompromisu rýchlejšieho nájdenia čo najlepšieho riešenia to je stratégia pravdepodobnostného kombinovania operátorov (s použitím aspoň jedného z operátorov REVERSE a GSR).

#### LITERATÚRA

- [1] W.J.Cook, *In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation*. Princeton University Press, 2011.
- [2] H.H. Hoos a T. Stutzle, "Stochastic Local Search Algorithms: An Overview" in *Handbook of Computational Intelligence*, Kacprzyk-Pedricz (Eds), Springer-Verlag, Berlin-Heidelberg, 2015, 1085-1105.
- [3] L. Sengupta, R. Mariescu-Istodor a F. Pasi, "Which Local Search Operator Works Best for the Open-loop TSP?," *Applied Sciences*, roč. 9, 2019, 24 strán.
- [4] E. Onder, M. Ozdemir a B.F. Yildirim, "Combinatorial optimization using artificial bee colony algorithm and particle swarm optimization supported genetic algorithm", *Kafkas University Journal of Economics and Administrative Science Faculty*, roč. 4, č. 6, 2013, 59-70.
- [5] N. Allahverdi, "Development a new mutation operator to solve the Traveling Salesman Problem by aid of genetic algorithms", *Expert Systems with Applications*, roč. 38, č. 3, 2011, 1313-1320.
- [6] TSPLIB - library of sample instances for the TSP. Dostupné: <http://elib.zib.de/pub/mp-testdata/tsp/tsplib/tsplib.html> (20.5.2021)
- [7] Z. Michalewicz a D.B.Fogel, *How to Solve It: Modern Heuristics*, 2nd ed. Springer-Verlag, Berlin-Heidelberg, 2010.
- [8] J. Frivaldský a M. Mach, "Porovnanie výberových stratégií horolezeckého algoritmu v rámci domény TSP" in *Electrical Engineering and Informatics XI*, Technical University Košice, 2020, 10-15.
- [9] J. Frivaldský, *Riešenie úlohy typu Problém obchodného cestujúceho pomocou prehľadavacích algoritmov*, Bakalárska práca, Fakulta elektrotechniky a informatiky Technická univerzita v Košiciach, 2020.



# Spracovanie EKG záznamov metódami strojového učenia

*Peter FOCKO, Ján GENČI*

Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická Univerzita v Košiciach, Slovenská republika

fockojr@gmail.com, jan.genci@tuke.sk

**Abstrakt** — Cieľom práce bolo rozšíriť existujúci systém pre spracovanie a vizualizáciu EKG záznamov o schopnosť kvantifikačného ohodnotenia známych príznakov srdcového infarktu v EKG. Výsledky kvantifikácie boli použité pre natrénovanie algoritmu strojového učenia s učiteľom, ktorý slúži ako automatizovaný detektor srdcového infarktu na základe EKG záznamu.

**Kľúčové slová** — analýza EKG, detekcia infarktu, spracovanie EKG, strojové učenie

## I. ÚVOD

Elektrokardiografické záznamy predstavujú informačný zdroj o aktuálnej činnosti a stave srdca pacienta. Pomocou prístroja nazývaného elektrokardiograf je možné merať výchylky hladiny elektrického potenciálu na rôznych miestach povrchu tela pacienta. Výchylky elektrického potenciálu na povrchu hrude a horných končatín ľudského tela sú spôsobené aktivitou srdcového svalového tkaniva. EKG záznam obsahuje namerané hodnoty elektrického potenciálu v čase z rôznych zvodov, respektíve elektród umiestnených na tele pacienta. Pre účely diagnostiky srdcových ochorení sú EKG záznamy primárny dátový zdroj.

Infarkt myokardu je porucha srdcovej činnosti, pri ktorej dochádza k výraznému zúženiu priechodnosti, alebo až upchatiu koronárnej artérie (srdcovej tepny). Upchatie koronárnej artérie môže už po niekoľkých minútach spôsobiť odumieranie svalových buniek, čo môže mať pre človeka až fatálne následky. Vo vyspelých krajinách patria infarkt myokardu a následné pretrvávajúce zdravotné komplikácie medzi najčastejšie príčiny úmrtia [1].

EKG záznamy sú v súčasnosti ešte stále štandardne manuálne analyzované kardiológmi. Využitie výpočtovej techniky pre spracovanie a analýzu EKG záznamov má potenciál pomôcť kardiológom v ich každodennej práci, čo v konečnom dôsledku môže mať pozitívny vplyv na kvalitu zdravotnej starostlivosti. Automatizovaná detekcia srdcových ochorení na základe zhotovených EKG záznamov je predmetom intenzívneho výskumu už niekoľko rokov. Funkčná implementácia automatického detektora srdcového infarktu môže v pomerne krátkom čase odhaliť pomerne veľké množstvo srdcových infarktov, prípadne poskytnúť kardiológom iný názor, respektíve upozornenie.

Predkladaná práca nadväzuje na projekt *Počítačové spracovanie EKG záznamov* [2]. Predmetom projektu bolo automatizované spracovanie EKG záznamov a všeobecná štatistická analýza najčastejšie sledovaných prvkov EKG záznamov, ako napríklad RR interval. Štatistické analýzy sa práci v rámci projektu zaoberali len časovými charakteristikami prvkov EKG záznamov, ako napríklad dĺžka, či premenlivosť dĺžky konkrétneho intervalu. Cieľom aktuálnej práce bolo navrhnúť a implementovať ďalšie, sofistikovanejšie štatistické analýzy EKG záznamov, ktoré sú cielené na kvantitatívne ohodnotenie známych príznakov infarktu myokardu. Motiváciou pre implementáciu cielených štatistických analýz je využitie ich výsledkov ako vstupné dáta pre tréning algoritmu strojového učenia. Cieľom výsledného predikčného modelu je rozlíšiť EKG záznamy zdravých pacientov od infarktových EKG záznamov s čo najväčšou úspešnosťou. Kvalita klasifikácia EKG záznamov bola experimentálne overená na vzorke verejne dostupných dát z portálu PhysioNet.

## II. ZDROJE DÁT

Pre účely tréningu a validácie predikčného modelu na detekciu infarktu bola potrebná dostatočne veľká vzorka dát, ktorá obsahuje infarktové EKG záznamy, ako aj normálne EKG záznamy bez symptómov srdcovej choroby. Okrem dostatočne veľkého počtu EKG záznamov

bolo potrebné, aby záznamy obsahovali aj vzorky signálov zo všetkých štandardných zvodov 12-zvodového elektrokardiografu: *I, II, III, aVR, aVF, aVL, V1, V2, V3, V4, V5* a *V6*, pretože prejavy infarktu myokardu na EKG sa môžu nachádzať v rozličných zvodoch. Dôležitá bola samozrejme aj kvalita digitálnych EKG záznamov a konzistentnosť ich formátu pre účely automatizovaného spracovania. Výber vhodného dátového zdroja pre túto prácu ovplyvňovala aj dostupnosť a kvalita metadát dostupných k EKG záznamom. V rámci metadát k EKG záznamom sú zaujímavé predovšetkým diagnostické závery kardiológov. Na základe vopred známych diagnostických informácií je možné vzorku dát rozdeliť na EKG záznamy zdravých pacientov a EKG záznamy s príznakmi srdcového infarktu, a následne použiť metódy strojového učenia s učiteľom (z angl. *supervised machine learning*) pre implementáciu automatizovaného detektora infarktu myokardu.

Všetky stanovené požiadavky na vhodný dátový zdroj pre túto prácu spĺňa databáza EKG záznamov PTB-XL [3], ktorá je verejne dostupná v rámci portálu PhysioNet. Táto databáza EKG záznamov obsahuje až 21837 rôznych EKG záznamov, ktoré boli zhotovené vyšetrením až 18885 rôznych pacientov. Všetky EKG záznamy v databáze PTB-XL sú reprezentované vo formáte MIT so vzorkovacou frekvenciou 500 Hz. Ku každému EKG záznamu v databáze PTB-XL je navyše dostupná aj jeho redukovaná verzia vo vzorkovacej frekvencii 100 Hz. Z hľadiska kompozície obsahuje databáza PTB-XL až 9528 normálnych EKG záznamov bez známok srdcových chorôb a až 5486 EKG záznamov, v ktorých možno pozorovať niektoré z príznakov infarktu myokardu. K osobitným záznamom databázy PTB-XL sú taktiež dostupné kvalitne spracované metadáta v štruktúrovanej forme. Metadáta sú reprezentované pomocou CSV súboru, kde každý riadok reprezentuje informácie o konkrétnom EKG zázname. Diagnostické informácie sú reprezentované ako pravdepodobnosť výskytu konkrétnej diagnózy.

### III. ARCHITEKTÚRA SYSTÉMU

Architektúru pôvodnej verzie systému možno považovať za monolitickú. Systém pozostával len z jedinej multifunkčnej webovej aplikácie, ktorá sa súbežne starala o spracovanie, vizualizáciu a analýzu EKG záznamov. Nakoľko pre aktuálny projekt bolo požiadavkou, aby bol výsledný systém efektívne horizontálne škálovateľný a nasaditeľný v klastrovom prostredí, architektúra systému bola zmenená na architektúru, založenú na mikroslužbách. Výsledný systém preto pozostáva z niekoľkých, navzájom kolaborujúcich mikroslužieb. Pre riadenie a nasadenie týchto mikroslužieb sa používa orchestračný systém Kubernetes, ktorý umožňuje spúšťať, škálovať a monitorovať mikroslužby v osobitných Docker kontajneroch. Pomocou systému Kubernetes je možné centrálné ovládať klaster výpočtových zariadení, do ktorého možno dynamicky pridávať ďalšie zariadenia s cieľom zvýšiť celkovú výpočtovú kapacitu nasadeného systému. Výsledné riešenie je adaptovateľné na aktuálne výkonnostné požiadavky pomocou horizontálneho škálovania, respektíve replikácie jednotlivých komponentov. Docker kontajnery navyše zjednodušujú prenositeľnosť jednotlivých mikroslužieb, nakoľko sa v kontajneroch nachádza aj predkonfigurované prostredie, špecifické pre konkrétny komponent. Nasledujúca schéma (Fig. 1) zachytáva jednotlivé mikroslužby v systéme a závislosti medzi nimi:

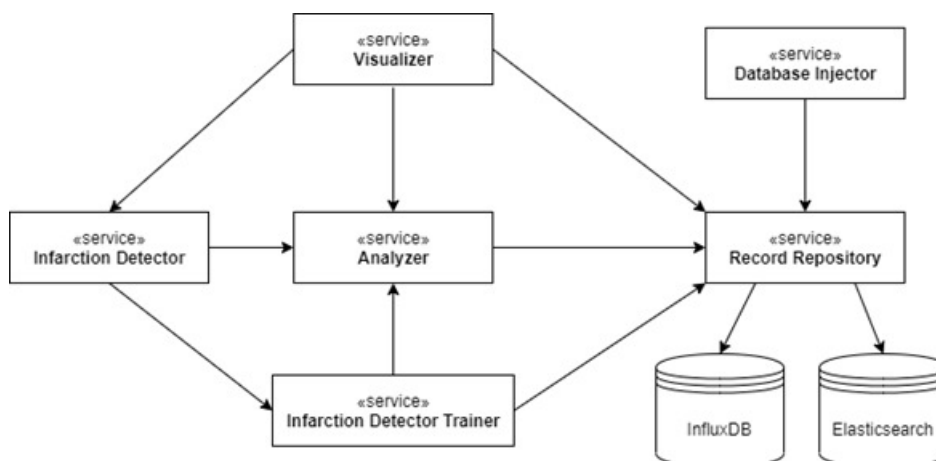


Fig. 1 Diagram závislostí medzi mikroslužbami v systéme.

#### A. Mikroslužba Record Repository

Mikroslužba *Record Repository* poskytuje HTTP aplikačné rozhranie pre ukladanie a získavanie EKG záznamov a k nim prislúchajúcim anotáciám a metadátam. Na pozadí používa repozitár EKG záznamov dva databázové systémy – InfluxDB a Elasticsearch. InfluxDB je NoSQL databázový systém, ktorý je špeciálne usporiadaný pre ukladanie časových radov, akými

sú aj vzorky signálov v EKG zázname. Elasticsearch je databázový systém na ukladanie a vyhľadávanie dokumentov vo formáte JSON. Elasticsearch sa v systéme používa pre ukladanie štruktúrovaných, alebo čiastočne štruktúrovaných metadát, prislúchajúcim k jednotlivým EKG záznamom. Úlohou repozitára EKG záznamov je zabezpečovať konzistenciu dát naprieč obidvoma databázovými systémami a taktiež slúži ako jediný prístupový bod pre ostatné mikroslužby na vstupno-výstupné operácie nad EKG dátami. Mikroslužbu *Record Repository* možno horizontálne škálovať.

#### B. Mikroslužba Visualizer

Úlohou mikroslužby *Visualizer* je poskytovať webové rozhranie systému, ktoré je určené pre koncového používateľa. Vizualizačná mikroslužba vznikla extrakciou vizualizačnej funkcionality z pôvodného systému do samostatnej mikroslužby. Táto mikroslužba umožňuje vo webovom prehliadači zobraziť konkrétny EKG záznam, výsledky štatistických analýz nad konkrétnym záznamom a výsledok detekcie infarktu vo zvolenom EKG zázname. Pre získavanie EKG dát využíva aplikačné rozhranie mikroslužby *Record Repository* a pre získavanie výsledkov štatistických analýz používa rozhranie mikroslužby *Analyzer*. Výsledky detekcie infarktu sú dostupné z aplikačného rozhrania mikroslužby *Infarction Detector*. Mikroslužba *Visualizer* je horizontálne škálovateľná.

#### C. Mikroslužba Analyzer

Mikroslužba *Analyzer* slúži na vykonávanie rôznych druhov štatistických analýz nad konkrétnym EKG záznamom. Táto mikroslužba vznikla extrakciou analytickej funkcionality z pôvodnej verzie systému. Pôvodné dva podporované druhy štatistických analýz – analýza *RR* intervalov a analýza intervalov srdcového cyklu boli doplnené o dva nové druhy štatistických analýz – analýza *ST* segmentov a analýza *QRS* komplexov. Analyzátor EKG záznamov, ako samostatná mikroslužba v systéme, je závislý len na repozitári EKG záznamov (mikroslužba *Record Repository*). Ostatným mikroslužbám ponúka analyzátor EKG záznamov výsledky štatistických analýz prostredníctvom aplikačného HTTP rozhrania. Mikroslužbu *Analyzer* možno horizontálne škálovať.

#### D. Mikroslužba Database Injector

Mikroslužba *Database Injector* sa v systéme používa na import celej databázy EKG záznamov z portálu PhysioNet. V rámci tejto mikroslužby bol implementovaný všeobecný algoritmus pre spracovanie a odoslanie všetkých EKG záznamov ľubovoľnej databázy EKG záznamov z portálu PhysioNet. Počas spracovania EKG záznamov sú k jednotlivým EKG záznamom vygenerované anotácie, ktoré označujú konkrétne prvky v zázname, ako napríklad začiatok, vrchol alebo koniec špecifickej vlny. Pre generovanie anotácií boli použité už existujúce nástroje *gqrs* [4] a *ecgpuwave* [5]. *Database Injector* umožňuje taktiež automatizovane spracovať metadáta, ktoré sú ku konkrétnym EKG záznamom dostupné. Formát a štruktúra metadát sa však naprieč rôznymi databázami z portálu PhysioNet líši. Mikroslužba *Database Injector* preto zahŕňa implementáciu dodatočných procesorov metadát, ktoré umožňujú transformovať metadáta z formátu špecifického pre konkrétnu databázu do jednotnej formy. Mikroslužba *Database Injector* je závislá jedine na repozitári EKG záznamov. Táto mikroslužba po ukončení svojej činnosti terminuje, spúšťa sa len v prípade potreby importu nových dát. Nakoľko niektoré databázy EKG záznamov z portálu PhysioNet sú pomerne obsiahle, *Database Injector* umožňuje paralelizovať spracovanie a importovanie dát na zvolenom počte procesorov.

#### E. Mikroslužba Infarction Detector Trainer

Úlohou mikroslužby *Infarction Detector Trainer* je zhromaždiť výsledky štatistických analýz pre všetky vybrané EKG záznamy a následne ich použiť pre natrénovanie predikčného modelu na detekciu infarktu myokardu. Pre načítanie zoznamu normálnych a infarktových EKG záznamov používa *Infarction Detector Trainer* aplikačné rozhranie repozitára EKG záznamov. Potom sa pre každý EKG záznam z vytvoreného zoznamu načítajú výsledky všetkých druhov štatistických analýz pomocou aplikačného rozhrania mikroslužby *Analyzer*. Zhromaždené dáta sa následne použijú pre natrénovanie algoritmu strojového učenia s učiteľom. Výsledný predikčný model je po natrénovaní exportovaný na dátový priestor, ktorý je zdieľaný s mikroslužbou *Infarction Detector*. V rámci mikroslužby *Infarction Detector Trainer* je taktiež možné uplatniť princípy dátového paralelizmu s cieľom skrátenia celkového času tréningu. Po exportovaní predikčného modelu táto mikroslužba terminuje.

#### F. Mikroslužba Infarction Detector

Mikroslužba *Infarction Detector* slúži na vykonanie detekcie infarktu vo zvolenom EKG zázname. Pre vykonanie detekcie sa využíva predikčný model, ktorý bol importovaný z dátového úložiska, ktoré je zdieľané so službou *Infarction Detector Trainer*. Detektor infarktu, ako mikroslužba v systéme je závislý na analyzátoch EKG záznamov, nakoľko výsledky štatistických

analýz sa používajú ako vstupné dáta pre vykonanie detekcie. Detektor srdcového infarktu poskytuje výsledky detekcie pomocou vlastného aplikačného HTTP rozhranie a je možné ho horizontálne škálovať.

#### IV. ŠTATISTICKÁ ANALÝZA EKG ZÁZNAMOV

Pôvodná verzia systému podporovala len dva druhy štatistických analýz – analýza *RR* intervalov a analýza intervalov srdcového cyklu. Tieto dva druhy analýz sa zaoberali len časovými charakteristikami sledovaných prvkov EKG, ako napríklad priemerná dĺžka intervalu alebo premenlivosť dĺžky konkrétneho intervalu. Pre účely kvantifikačného ohodnotenia príznakov infarktu myokardu je však potrebné zaoberať sa aj tvarovými vlastnosťami EKG krivky v osobitných zvodoch EKG. Do systému bola teda pridaná podpora pre vykonanie štatistickej analýzy *ST* segmentov a *QRS* komplexov nad konkrétnym EKG záznamom. Hlavnou motiváciou pre implementáciu týchto nových druhov štatistických analýz EKG záznamov je použitie ich výsledkov pre natréňovanie predikčného modelu na detekciu infarktu. Úspešné vykonanie štatistickej analýzy *ST* segmentov a *QRS* komplexov je závislé od prítomnosti vopred vygenerovaných anotácií pomocou programu *ecgpuwave* [5]. Výsledný systém umožňuje vizualizáciu výsledkov všetkých druhov štatistických analýz v podobe interaktívnych grafov v rámci webového používateľského rozhrania.

##### A. Štatistická analýza *ST* segmentov

Primárne príznaky infarktu myokardu sa v rámci EKG nachádzajú v oblasti *ST* segmentu (Fig. 2). Podľa príznakov srdcového infarktu je možné infarkty rozdeliť na dve skupiny - *STEMI* a *NSTEMI*. Infarkty typu *STEMI* sa vyznačujú predovšetkým vyvýšenou hladinou *ST* segmentu [6]. Výrazná elevácia *ST* segmentu môže dokonca spôsobiť tvarovú deformáciu nasledujúcej *T* vlny alebo predchádzajúceho *S* kmitu. Najčastejším príznakom infarktov typu *NSTEMI* je naopak depresia [7], respektíve zníženie hladiny *ST* segmentu. Infarkty typu *NSTEMI* sa vyznačujú aj prítomnosťou ischemických *T* vln. Ischemické *T* vlny majú v konkrétnom EKG zvode opačnú orientáciu ako *R* kmity, teda ak má *R* kmit v EKG zvode pozitívny tvar (smerom nahor), tak má ischemická *T* vlna v tom istom zvode negatívny tvar (smerom nadol).

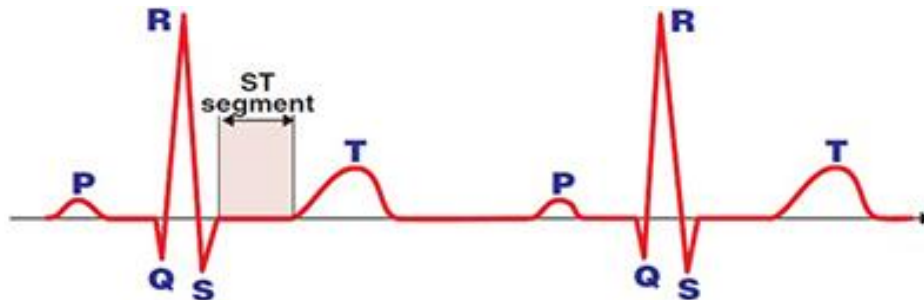


Fig. 2 Vizualizácia jednotlivých vln a kmitov v EKG [8].

Pre účely implementácie automatického detektora infarktu myokardu bolo potrebné navrhnuť vhodné metriky, ktoré popisujú uvedené príznaky srdcového infarktu v oblasti *ST* segmentu a *T* vlny. Ako základná metrika bola zvolená priemerná vertikálna odchýlka hladiny *ST* segmentu od priemernej hladiny izoelektrickej čiary (úsek EKG záznamu kde sa žiadna vlna alebo kmit nenachádza). Ďalej bolo potrebné opísať tvar *ST* segmentu. Zvolené boli dve metriky – uhol stúpania lineárnej funkcie aproximovanej k *ST* segmentu a priemerná zmena hladiny *ST* segmentu. Pre zachovanie istej formy relativity k ostatným dôležitým prvkom EKG ako napríklad *T* vlna a *QRS* komplex boli navrhnuté ďalšie dve metriky – relatívna odchýlka *ST* segmentu v porovnaní s maximálnou odchýlkou *QRS* komplexu a relatívna odchýlka *ST* segmentu v porovnaní s maximálnou odchýlkou *T* vlny. Pre účely charakterizácie tvaru *T* vlny boli pridané aj metriky – priemerná odchýlka hladiny *T* vlny od izoelektrickej čiary a relatívna veľkosť *T* vlny oproti veľkosti *QRS* komplexu. V neposlednom rade sa počíta aj dĺžka *ST* segmentov a *T* vlny. Všetky uvedené metriky sa počítajú pre každý zvod, respektíve signál v EKG zázname osobitne. Medzi výsledkami štatistickej analýzy *ST* segmentov možno nájsť aj sumarizačné štatistiky, v rámci ktorých sa uvádza aritmetický priemer a štandardná odchýlka hodnôt každej použitej metriky pre konkrétny zvod EKG záznamu.

##### B. Štatistická analýza *QRS* komplexov

V rámci oblasti *QRS* komplexu (Fig. 2) možno pozorovať sekundárne príznaky infarktu myokardu. Typickým príznakom srdcového infarktu v oblasti *QRS* komplexu je napríklad patologický kmit *Q* [8]. Patologický kmit *Q* (Fig. 3) je označenie pre nadmerne zväčšený kmit *Q*. V oblasti *QRS* komplexu možno pozorovať aj iné deformácie, ktoré sú spôsobené výrazným

porušením hladiny *ST* segmentu, ako napríklad optický zánik kmitu *S* (Fig. 3) alebo *Pardeeho vlny* [6], respektíve kompletne prekrytie *QRS* komplexu zvýšenou hladinou *ST* segmentu.

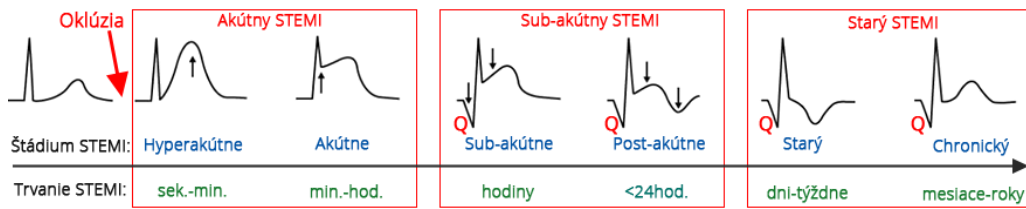


Fig. 3 Vizualizácia elevácie *ST* segmentov a patologických *Q* kmitov v EKG [8]

Uvedené symptómy v rámci *QRS* komplexu bolo potrebné charakterizovať vhodnými metrikami. Súčasťou štatistickej analýzy *QRS* komplexov je výpočet veľkostí jednotlivých kmitov *Q*, *R*, *S*. Za veľkosť konkrétneho kmitu sa považuje odchýlka jeho amplitúdy od priemernej hladiny izoelektrickej čiary. Pre účely vyjadrenia relativity medzi jednotlivými kmitmi *QRS* komplexu boli navrhnuté dve metriky – vertikálny rozdiel amplitúdy kmitu *R* a špičky kmitu *Q*, a vertikálny rozdiel amplitúdy kmitu *R* a špičky kmitu *S*. Relativitu veľkosti *QRS* komplexu zachytáva aj metrika, ktorá vyjadruje relatívnu veľkosť vlny *P* oproti veľkosti *QRS* komplexu (vypočítaná ako maximálna odchýlka od izoelektrickej čiary). V neposlednom rade sa počíta aj dĺžka *QRS* komplexu a *P* vlny. Všetky uvedené metriky je navyše potrebné vypočítať pre každý EKG zvod samostatne. Štatistická analýza *QRS* komplexov zahŕňa aj sumarizačné štatistiky, kde je vyčíslený aritmetický priemer a štandardná odchýlka hodnôt každej uvedenej metriky pre špecifický zvod, respektíve signál EKG záznamu.

## V. TRÉNOVANIE PREDIKČNÉHO MODELU

Hlavným cieľom práce bolo natrénovanie predikčného modelu, ktorý bude schopný rozlíšiť EKG záznamy zdravých pacientov a EKG záznamy pacientov s infarktom. Nakoľko boli v rámci databázy PTB-XL [3] dostupné EKG záznamy už aj s diagnostickými informáciami od kardiológov, použité boli metodológie strojového učenia s učiteľom. Z typologického hľadiska predstavuje implementovaný predikčný model binárny klasifikátor, nakoľko umožňuje klasifikovať EKG záznamy do dvoch klasifikačných tried – normálne a infarktové. Vstupnými dátami pre trénovanie algoritmov strojového učenia boli sumarizačné štatistiky vykonaných štatistických analýz a vopred známe diagnostické informácie.

Dokopy sa pre účely trénovaní a validácie predikčného modelu použilo 7831 EKG záznamov z databázy PTB-XL, z ktorých 5633 patrilo zdravým pacientom a 2198 patrilo pacientom s infarktom. EKG záznamy zdravých pacientov predstavovali záznamy z databázy PTB-XL, ktoré mali podľa diagnostických informácií v metadátoch stanovenú pravdepodobnosť normálneho EKG 100%. Za infarktové EKG záznamy sa považovali také záznamy, pri ktorých bola uvedená pravdepodobnosť výskytu infarktu myokardu taktiež 100%. Zo vstupnej dátovej sady boli vyradené EKG záznamy, ktoré mali v metadátoch zmienku o prípadných technických problémoch v niektorých zvodoch EKG. Použitie problémových EKG záznamov by mohlo zapríčiniť nevyspytateľné výsledky. V rámci hľadania a konfigurácie vhodného predikčného modelu sa 80% EKG záznamov používalo na trénovanie a 20% na testovanie.

V kontexte použitých metrik na ohodnotenie kvality predikčného modelu sa používala celková úspešnosť predikcie, respektíve percento správne klasifikovaných záznamov. Nakoľko je ale pomerové zastúpenie klasifikačných tried v trénovej a testovacej časti dát nevyrovnané, bolo potrebné použiť aj iné metriky. Nakoľko prioritou predikčného modelu je pravdivo klasifikovať čo najväčšie percento infarktových EKG záznamov, merala sa aj percentuálna miera pravdivo pozitívnych prípadov vzhľadom k všetkým pozitívnym (infarktovým) prípadom. Ďalšou použitou metrikou bola metrika F1 skóre, pretože táto metrika je vhodná pre nerovnomerné zastúpenie klasifikačných tried v trénovej a testovacej časti vstupnej dátovej sady.

V rámci hľadania vhodného predikčného modelu sa experimentovalo s viacerými algoritmi strojového učenia z knižnice scikit-learn [9]. Jednotlivé algoritmy je možné konfigurovať pomocou tzv. hyperparametrov. Pre odskúšanie rôznych kombinácií hodnôt hyperparametrov pre daný algoritmus bol použitý GridSearchCV [9], ktorý pre ohodnotenie kvality danej konfigurácie modelu používa metódu krížovej validácie na trénovej časti vstupných dát. Kombinácia hodnôt hyperparametrov s najlepšimi výsledkami pre zvolený algoritmus bola následne nezávisle otestovaná na vopred vymedzenej testovacej časti dát. V tabuľke Tab. 1 sú na základe stanovených metrik uvedené najlepšie dosiahnuté výsledky pre konkrétne odskúšané algoritmy:

Dosiahnuté výsledky odskúšaných algoritmov

Algoritmus	F1 skóre	Presnosť	Podiel pravdivo pozitívnych predikcií
Support Vector Classifier	0,91	95,53%	92,05%
Random Forest Classifier	0,91	95,66%	88,41%
Ada Boost Classifier	0,93	95,6%	89,55%
Gradient Boosting Classifier	0,94	97,06%	93,41%
Linear Support Vector Classifier	0,88	93,43%	89,55%

Ako je možné v tabuľke Tab. 1 vidieť, najlepšie výsledky predikcie boli dosiahnuté pomocou algoritmu *Gradient Boosting Classifier*, ktorý dokázal správne klasifikovať vyše 97% EKG záznamov a správne identifikovať až vyše 93% zo všetkých skutočne infarktových záznamov. Hodnoty hyperparametrov algoritmu *Gradient Boosting Classifier* boli nasledovné: *criterion = mse*, *learning\_rate = 0.4*, *loss = deviance* a *n\_estimators = 400*. Predikčný model s touto konfiguráciou bol následne natrénovaný na celej vzorke tréningových dát a použitý v konečnej verzii systému na vykonanie detekcie infarktu myokardu v konkrétnom EKG zázname.

#### ZÁVER

Výsledkom práce je rozšírenie existujúceho systému na automatizované spracovanie a vizualizáciu EKG záznamov. Pôvodný monolitický systém bol dekomponovaný na niekoľko navzájom kolaborujúcich mikroslužieb, ktoré sú horizontálne škálovateľné a prispôbené pre nasadenie v klastri. Do systému bola pridaná podpora pre vykonanie dvoch nových druhov štatistických analýz nad EKG záznamami – štatistická analýza *ST* segmentov a štatistická analýza *QRS* komplexov. Výsledky štatistických analýz boli použité pre natrénovanie predikčného modelu, ktorý dokáže odlíšiť normálne EKG záznamy od infarktových EKG záznamov s pravdepodobnosťou pravdivej predikcie vyše 97%. Konečný používateľ systému môžu so systémom interagovať pomocou webového používateľského rozhrania, ktoré umožňuje v podobe interaktívnych grafov zobrazit' EKG záznam, výsledky predspracovania, výsledky vykonaných štatistických analýz a tiež aj výsledky detekcie infarktu. Konečná verzia systému je nasadená a dostupná verejnosti na adrese <https://www.ecgstack.com>.

#### POUŽITÁ LITERATÚRA

- [1] NEHAJ F. - SOKOL J. - Kovář F. - KUBAŠKOVÁ M. - MOKÁŇ M. - MOKÁŇ M.: Infarkt myokardu – rekapitulácia poznatkov. Unilabs. In *inVitro*. November 2017, roč. 3.
- [2] FOCKO P.: Počítačové spracovanie EKG záznamov: Bakalárska práca. Košice: TU FEI, 2019. 64s.
- [3] WAGNER P. - STRODTHOFF N. - BOUSSELJOT R. - KREISELER D. - LUNZE F. I. - SAMEK W. - SCHAEFFTER T.: PTB-XL, a large publicly available electrocardiography dataset. In *Scientific Data*. Máj 2020, roč. 7, s. 154.
- [4] MOODY G. B.: *gqrs*, *gqpost* - *QRS detector and post-processor* [online]. MIT Laboratory for Computational Physiology. PhysioNet. Marec 2019 [cit. 2021-05-11]. Dostupné na internete: <<https://www.physionet.org/physiotools/wag/gqrs-1.htm>>.
- [5] LAGUNA P. - JAN R. - BOGATELL E. - ANGLADA D. V.: *ecgpuwave - QRS detector and waveform limit locator* [online]. MIT Laboratory for Computational Physiology. PhysioNet. Marec 2019 [cit. 2021-05-11]. Dostupné na internete: <<https://physionet.org/physiotools/wag/ecgpuw-1.htm>>.
- [6] BLAHÚT P.: *STEMI infarkt* [online]. TECHMED. TECHMED. November 2017 [cit. 2021-05-11]. Dostupné na internete: <<https://www.techmed.sk/infarkt-myokardu-stemi>>.
- [7] BLAHÚT P.: *NSTEMI infarkt a Nestabilná angina pectoris* [online]. TECHMED. TECHMED. November 2017 [cit. 2021-05-11]. Dostupné na internete: <<https://www.techmed.sk/nstemi-infarkt-myokardu-a-instabilna-angina-pectoris>>.
- [8] HOUGHTON, ANDREW R.: Making sense of the ECG : a hands-on guide. London: CRC Press, 2019.
- [9] PEDREGOSA F. - VAROQUAUX G. - GRAMFORT A. - MICHEL V. - THIRION B. - GRISEL O. - BLONDEL M. - PRETTENHOFER P. - WEISS R. - DUBOURG V. - VANDERPLAS J. - PASSOS A. - COURNAPEAU D. - BRUCHER M. - PERROT M. - DUCHESNAY E.: Scikit-learn: Machine Learning in Python. In *Journal of Machine Learning Research*. 2011, roč. 12, s. 2825-2830.

# Softvér v programovacom prostredí LabView určený na sledovanie pohybujúcich sa osôb v reálnom čase

<sup>1</sup>Tamás PORTELEKY, <sup>2</sup>Dušan KOCUR, <sup>3</sup>Jana FORTES

<sup>1</sup>K-Mlab organizačná zložka Ilmsens GmbH, Košice, Slovensko

<sup>2,3</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky  
a informatiky, Technická univerzita v Košiciach

<sup>1</sup>Tamas.Porteley@ilmsens.com, <sup>2</sup>Dusan.Kocur@tuke.sk, <sup>3</sup>Jana.Fortes@tuke.sk

**Abstrakt** – V nasledujúcom príspevku je opísaný návrh programu UWB-PerLoc-LAB, ktorý je určený na detekciu, lokalizáciu a sledovanie cieľov v reálnom čase. UWB-PerLoc-LAB je softvérové riešenie vytvorené v programovacom prostredí LabView, v ktorom je implementovaná komunikácia medzi počítačom a radarom, algoritmy spracovania radarových signálov a grafické zobrazovanie výsledkov spracovania signálov. Okrem „on-line“ verzie programu pracujúcej v reálnom čase, bola vytvorená aj „off-line“ verzia softvéru UWB-PerLoc-LAB, ktorá umožňuje prácu so softvérom aj bez priameho pripojenia k radaru. Obe verzie boli navrhnuté tak, aby ich bolo možné používať ako samostatné aplikácie bez potreby nainštalovaného programovacieho prostredia LabView. V závere príspevku uvádzame príklady využitia softvéru UWB-PerLoc-LAB v praxi a námety na jeho rozšírenie.

**Kľúčové slová** – sledovanie pohybujúcich sa osôb, softvér v programovacom prostredí LabView, spracovanie radarových signálov v reálnom čase, UWB radar

## I. ÚVOD

UWB (z angl. *Ultra WideBand*) senzorové systémy sú v súčasnosti predmetom intenzívneho výskumu. Táto technológia môže byť okrem iného využitá aj na detekciu, lokalizáciu a sledovanie osôb, ktoré má široké využitie v praxi. Napríklad v prípade živelných pohrôm je detekcia a lokalizácia osôb uväznených pod sutinami veľmi náročnou úlohou. Taktiež veľký počet osôb a ich záchrancov zostáva zakliesnených v budovách plných dymu pri nulovej optickej viditeľnosti. V takýchto prípadoch by na vyhľadanie osôb mohol byť s výhodou použitý UWB radar (senzor) vďaka svojej schopnosti detegovať osoby za nekovovou prekážkou [1].

Pre vývoj aplikácií UWB senzorov na detekciu a lokalizáciu osôb je užitočné disponovať softvérom pracujúcim v reálnom čase. Zámerom bakalárskej práce [2] bolo preto navrhnuť softvér určený na spracovanie radarových signálov s cieľom detegovať, lokalizovať a sledovať pohybujúce sa ciele v reálnom čase. Keďže pri jeho návrhu bolo potrebné používať už existujúce kódy na spracovanie radarových signálov, ktoré boli vytvorené v programovacom jazyku C, bolo nevyhnutné, aby vývojové prostredie podporovalo použitie dynamických knižníc. Ďalej bolo potrebné, aby vývojové prostredie poskytovalo nástroje na vytvorenie grafického používateľského rozhrania. Nakoľko išlo o spracovanie signálov v reálnom čase, požiadavka na dostatočnú rýchlosť spracovania bola tiež veľmi dôležitá. Programovacie prostredie LabVIEW (z angl. *Laboratory Virtual Instrument Engineering Workbench*) tieto požiadavky pri návrhu programov dovoľovalo naplniť a práve preto bolo LabVIEW vhodné na vytvorenie nového softvéru nazvaného UWB-PerLoc-LAB [3]. Na softvér UWB-PerLoc-LAB boli zadané nasledovné požiadavky: používateľsky prívetivý dizajn, grafické zobrazovanie vybraných veličín vznikajúcich v procese spracovania radarových signálov a jednoduchá manipulácia s parametrami metód spracovania signálov. Okrem toho, UWB-PerLoc-LAB mal byť navrhnutý tak, aby bol použiteľný ako samostatná aplikácia [2].

Na vytvorenie takéhoto softvéru bolo potrebné riešiť tri úlohy. Prvou úlohou bola implementácia aplikačného programovacieho rozhrania pre komunikáciu medzi počítačom a radarom. Druhou úlohu predstavovala implementácia algoritmov spracovania radarových signálov a treťou úlohou bol návrh grafického používateľského rozhrania so žiadanými vlastnosťami ako je napr. používateľsky prívetivý dizajn. Riešenie týchto úloh je opísané v nasledujúcich častiach tohto príspevku.

## II. ZÁKLADNÁ KONCEPCIA KOMUNIKÁCIE MEDZI POČÍTAČOM A RADAROM

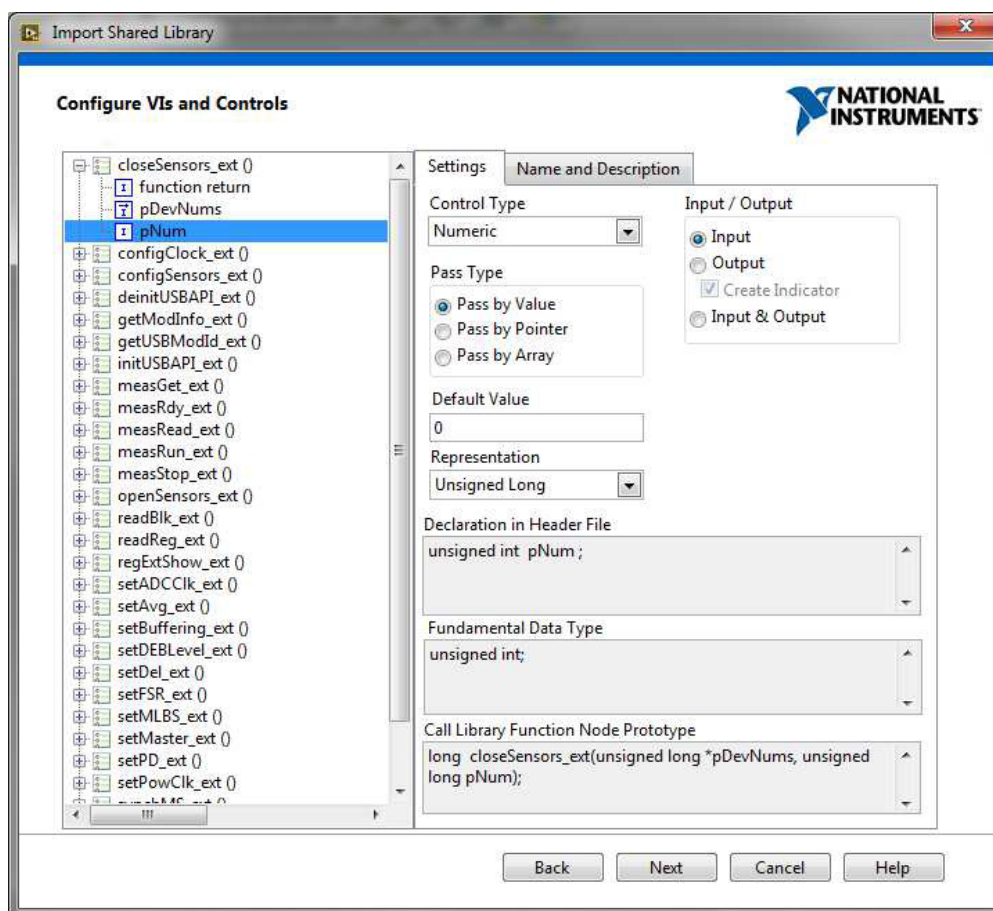
Program UWB-PerLoc-LAB pracuje s UWB radarom (senzorom) m:explore, ktorý emituje M-postupnosť 12-tého rádu. Tento radar bol vyvinutý spoločnosťou Ilmsens GmbH [4]. K radaru je dodávané aj aplikačné programové rozhranie (API – z angl. *Application Programming Interface*), ktoré zabezpečuje komunikáciu medzi používateľskou aplikáciou a radarom. Rozhranie API obsahuje súbor funkcií, ktoré umožňujú predovšetkým vykonávať nasledujúce úlohy:

- detekcia senzorov pripojených k počítaču,
- otvorenie a zatvorenie komunikáciu so senzorom,
- nastavenie a konfigurácia niektorých parametrov senzora,
- spustenie a zastavenie merania,
- prenos nameraných údajov do používateľskej aplikácie.

Napriek tomu, že tieto funkcie boli napísané v jazyku C, API je použiteľné v každom programovacom prostredí, v ktorom je možné volať funkcie z dynamicky pripájaných knižníc (DLL – z angl. *Dynamic Link Library*). Dokumentácia API rozhrania obsahuje opis všetkých knižničných funkcií a všetky informácie, ktoré je potrebné vedieť pri volaní týchto funkcií, t.j. vstupné a výstupné parametre funkcie, návratové hodnoty a ich dátové typy [5].

### A. Implementácia API rozhrania v prostredí LabVIEW

Rozhranie API je vlastne knižnica skompilovaných funkcií, ktoré sú uložené v súbore typu DLL. Keďže programovacie prostredie LabVIEW umožňuje implementáciu dynamických knižníc, aplikácia rozhrania API v LabVIEW je pomerne jednoduchá. Pri importovaní dynamickej knižnice do LabVIEW je potrebné mať k dispozícii aj hlavičkový súbor, ktorý obsahuje prototypy (deklarácie) všetkých funkcií tejto knižnice. Dôležitým krokom pri importovaní dynamickej knižnice je správne nastavenie dátových typov vstupných a výstupných parametrov funkcií (Obr. 1). Výsledkom procedúry importovania dynamickej knižnice je tzv. *subvi* (podprogram), ktorý sa vygeneruje pre každú jednu funkciu tejto knižnice. Použitím týchto podprogramov v blokovom diagrame je možné zavolať príslušnú funkciu z knižnice.



Obr. 1 Nastavenie parametrov funkcií pri importovaní dynamickej knižnice do LabView.



### III. SPRACOVANIE UWB RADAROVÝCH SIGNÁLOV V PROGRAME UWB-PERLOC-LAB

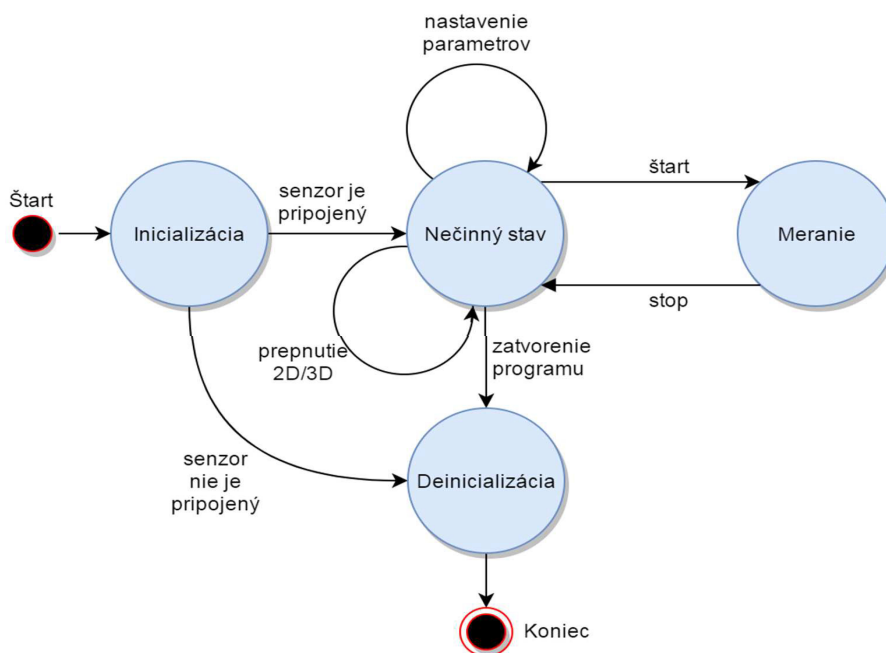
Prvým a nevyhnutným krokom spracovania radarových signálov je výpočet impulzovej odpovede vyšetřovaného priestoru. Impulzová odpoveď je získaná tak, že sa vypočíta vzájomná korelačná funkcia signálu prijatého radarom a vysielanej M-postupnosti. V programe UWB-PerLoc-LAB je tento výpočet realizovaný použitím rýchlej Fourierovej transformácie. Výsledkom výpočtu je impulzová odpoveď vyšetřovaného priestoru, ktorá v takomto formáte už môže byť spracovaná podľa procedúry spracovania UWB radarových signálov. Algoritmy tejto procedúry, ktoré sú výsledkom výskumu na Katedre elektroniky a multimediálnych telekomunikácií Technickej univerzity v Košiciach, sú napísané v jazyku C [1], [6]. Proces implementácie týchto funkcií v prostredí LabView je rovnaký, ako bol opísaný v predchádzajúcej časti tohto príspevku. Niektoré jednoduchšie metódy ako nastavenie času nula a odčítanie pozadia sú naprogramované priamo v prostredí LabView.

Celá procedúra spracovania UWB radarových signálov vrátane výpočtu impulzovej odpovede je implementovaná v podprograme „data\_processing“. Vstupom podprogramu je prijatý signál získaný z radaru a klaster (štruktúra), ktorý obsahuje všetky parametre, ktoré sú potrebné k aplikácii jednotlivých metód. Výstupom podprogramu je tiež klaster, ktorý obsahuje výstupy z každej etapy spracovania dát. Tieto výstupy sú potom použité na grafické zobrazovanie výsledkov z jednotlivých fáz spracovania signálov. Aby bolo možné v budúcnosti čo najjednoduchšie implementovať do programu UWB-PerLoc-LAB nové metódy, je spracovanie dát realizované na báze aplikácie princípu konečného automatu (z angl. *state machine pattern*), kde jednotlivé stavy automatu zodpovedajú jednotlivým metódam [7].

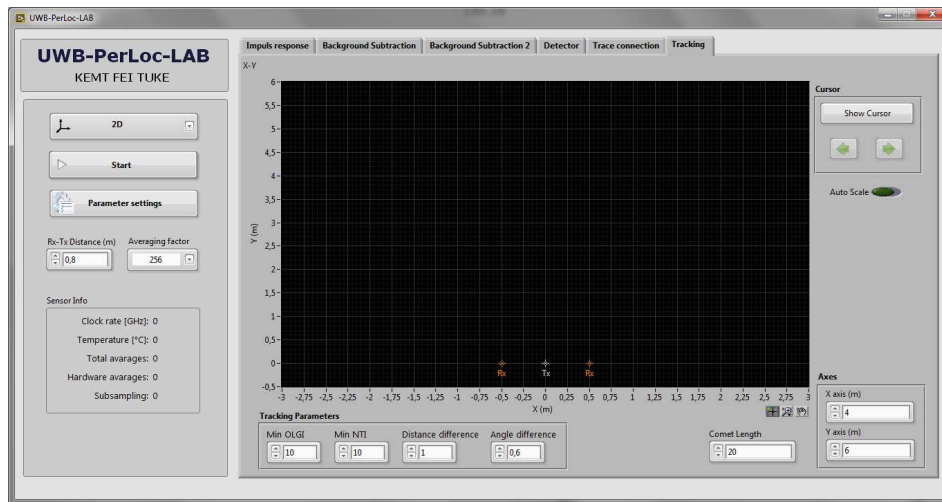
#### A. Konceptia programu UWB-PerLoc-LAB

Pri návrhu programu UWB-PerLoc-LAB bolo cieľom aj to, aby program bol použiteľný ako samostatná aplikácia. To má tú výhodu, že používateľ nemusí mať nainštalované na svojom počítači vývojové prostredie LabView, stačí nainštalovať balík, ktorý obsahuje tzv. *run-time engine* umožňujúci spustiť program UWB-PerLoc-LAB. Ďalšou výhodou je to, že používateľ nemá prístup k „zdrojovému kódu“.

Po implementovaní API rozhrania radaru m:explore a procedúry spracovania radarových signálov do programovacieho prostredia LabView bolo potrebné nájsť taký návrhový vzor, ktorý umožňuje, aby bol program použiteľný ako samostatná aplikácia. Vhodným riešením tohto problému bola aplikácia princípu konečného automatu a udalosťami riadeného programovania (z angl. *event driven programming*) [7], [8]. Stavový diagram programu UWB-PerLoc-LAB využívajúci túto koncepciu je uvedený na Obr. 2.



Obr. 2 Stavový diagram programu UWB-PerLoc-LAB.



Obr. 3 Používateľské rozhranie programu UWB-PerLoc-LAB.

#### IV. POUŽÍVATEĽSKÉ ROZHRAVIE

Na Obr. 3 je uvedené používateľské rozhranie programu UWB-PerLoc-LAB. Na ľavej strane sú umiestnené ovládacie prvky programu, pomocou ktorých je možné prepnúť medzi 2D a 3D režimom, spustiť meranie, nastaviť parametre metód spracovania radarových signálov, nastaviť vzdialenosť medzi vysielačou a prijímačou anténou radarového systému a nastaviť spriemerňovací faktor senzora. Spriemerňovací faktor hovorí o tom, koľko periód prijatého signálu má byť spriemerňovaných. Čím je táto hodnota vyššia, tým bude väčší odstup signálu od šumu, avšak rýchlosť merania bude menšia. V tejto sekcii sú tiež zobrazené základné informácie senzora ako napr. taktovacia frekvencia, teplota senzora atď.

Keďže procedúra spracovania UWB radarových signálov pozostáva z niekoľkých fáz a grafické výstupy z každej fázy obsahujú užitočné informácie pre operátora o ciele, je potrebné zobrazovať výstupy z každej fázy spracovania dát. Zobrazovanie týchto výstupov súčasne na štandardnom monitore nie je možné kvôli nedostatku miesta. Nehovoriac o tom, že by to spomalilo program, lebo by musel aktualizovať každý graf. V programovacom prostredí LabView možno vyriešiť tento problém pomocou nástroja „tab control“. To umožňuje vytvoriť viac strán, medzi ktorými možno prepnúť kliknutím na horné záložky. Na jednotlivých stránkach sú zobrazené výstupy z jednotlivých fáz spracovania dát a operátor môže vybrať, ktorý graf (grafy) chce vidieť, pričom je aktualizovaný vždy len ten graf, ktorý je umiestnený na vybranej stránke.

Program UWB-PerLoc-LAB okrem zobrazovania cieľov ponúka aj iné užitočné funkcie. Napr. pomocou tlačidla „Show cursor“ je možné zapnúť kurzor, ktorý sleduje pohyb cieľa. Ak v monitorovanom priestore je prítomných viac cieľov, pomocou zelenej šípky je možné vybrať, ktorý cieľ má byť sledovaný kurzorom. Kliknutím na tlačidlo „Hide cursor“ môžeme vypnúť túto funkciu. Ďalšou funkciou je „Auto scale“. Ak je zapnutá, rozsah zobrazovanej oblasti sa automaticky zväčšuje, ak sa cieľ vzdaluje. Takto sa nemôže stať, že sa cieľ bude pohybovať mimo zobrazovanej oblasti. Nastaviť rozsah zobrazovanej oblasti môžeme aj manuálne, pomocou numerických ovládačov, ktoré sú umiestnené v pravom dolnom rohu okna. Aby operátor videl trajektóriu sledovaného cieľa, sú zobrazené aj pozície cieľa odhadnuté v predchádzajúcich časových okamihoch pozorovania, pričom počet zobrazených predchádzajúcich pozícií možno nastaviť pomocou numerického ovládača „Comet length“.

Jednotlivé algoritmy spracovania UWB radarových signálov majú parametre, pomocou ktorých ich možno „ladiť“. V programe UWB-PerLoc-LAB na nastavenie týchto parametrov sú dve možnosti. Prvá možnosť je nastavenie parametrov počas merania. Na jednotlivých stranách „tab control“ sú umiestnené numerické ovládače, pomocou ktorých možno nastavovať parametre príslušnej metódy spracovania dát. V tomto prípade majú zmeny parametrov okamžitý efekt na algoritmy spracovania dát. Druhá možnosť je nastavenie parametrov, keď je meranie zastavené. Kliknutím na tlačidlo „Parameter settings“, ktoré je umiestnené na ľavej strane hlavného okna, sa objaví okno, v ktorom je možné nastavovať parametre všetkých metód spracovania dát. V tomto okne je dokonca možnosť uložiť parametre. Kliknutím na tlačidlo „Save“ sa objaví okno, v ktorom možno uložiť aktuálne nastavené parametre pod nejakým názvom, ku ktorým sa automaticky pripája aktuálny čas a dátum. Uložené parametre potom môžeme načítať pomocou tlačidla „Load“. Ďalšou možnosťou je resetovanie parametrov na predvolené hodnoty kliknutím na tlačidlo „Reset to default“.

## V. OFF-LINE VERZIA PROGRAMU UWB-PERLOC-LAB

Verzia opísaného programu UWB-PerLoc-LAB pracuje v reálnom čase. To znamená, že program kontinuálne načítava namerané údaje zo senzora, tieto údaje hneď spracováva a zobrazuje výsledky v reálnom čase. V takomto prípade ide o „on-line“ spracovanie dát. V prípade „off-line“ verzie programu UWB-PerLoc-LAB sa dáta nenačítavajú priamo zo senzora, ale zo súboru, ktorý obsahuje vopred namerané údaje. Najväčšou výhodou „off-line“ verzie programu je to, že počítač nemusí byť pripojený k senzoru, aby bolo možné otestovať algoritmy na spracovanie UWB radarových signálov. Takýmto spôsobom je možné program vylepšovať aj vtedy, keď UWB senzor nie je k dispozícii. Vykonané vylepšenia v „off-line“ verzii potom je možné implementovať do „on-line“ verzie programu.

Programom UWB-PerLoc-LAB je možné detegovať a lokalizovať pohybujúce sa ciele v 2D a 3D priestore. „Off-line“ verzia programu bola otestovaná v 2D a aj v 3D režime spracovaním experimentálne získaných údajov pre rôzne scenáre. „On-line“ verzia programu bola otestovaná len v 2D priestore, avšak program je pripravený na lokalizáciu pohybujúcich sa cieľov v reálnom čase aj v 3D priestore [2].

„On-line“ a „off-line“ verzie programu UWB-PerLoc-LAB sú navrhnuté tak, aby ich bolo možné používať ako samostatné aplikácie. V prílohe bakalárskej práce [2] sa nachádza aj inštallačný balík pre obidve verzie programu. Najväčšou výhodou používania týchto aplikácií je to, že inštallačný balík obsahuje aj LabView *Run-Time Engine*. Z tohto dôvodu na spustenie programu UWB-PerLoc-LAB nie je potrebné mať nainštalované programovacie prostredie LabView.

## VI. ZÁVER

Program UWB-PerLoc-LAB je v súčasnosti používaný na Katedre elektroniky a multimediálnych telekomunikácií na vývoj aplikácií UWB senzorov, ktorých účelom je lokalizovať osoby v reálnom čase. Ďalšou možnosťou praktického využitia je zakomponovanie tohto programového prostriedku do vzdelávacieho procesu na Technickej univerzite v Košiciach. Softvér UWB-PerLoc-LAB je vhodný aj na demonštračné účely pre potenciálnych zákazníkov.

Opísaný softvér je pripravený aj na lokalizáciu pohybujúcich sa cieľov v 3D priestore, avšak v reálnom čase ešte nebol testovaný. Preto bude v rámci ďalšieho vývoja tohto softvéru kladený dôraz aj na túto funkcionálnosť. Medzi ďalšie možné kroky pokračovania v práci na vývoji programu UWB-PerLoc-LAB patrí napr. implementácia nových metód spracovania radarových signálov a vylepšovanie grafického zobrazovania cieľov (hlavne v 3D režime).

## POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe projektu č. APVV-18-0373 a vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a Slovenskej akadémie vied (VEGA) na základe projektu č. 1/0584/20.

## LITERATÚRA

- [1] Kocur, D. et al., *Detekcia, lokalizácia a sledovanie osôb UWB radarmi krátkého dosahu*. Technická univerzita v Košiciach, Košice, Slovensko, 2016. ISBN 987-80-553-2648-1
- [2] Porteleky, T., *UWB-PerLoc-LAB: Softvér v programovacom prostredí LabView určený na vývoj aplikácií UWB senzorov na lokalizáciu osôb*. Bakalárska práca, Technická univerzita v Košiciach, Košice, Slovensko, 2017.
- [3] Benefits of Programming Graphically in NI LabVIEW. [online] [citované 2021]. Dostupný na internete: < <http://www.ni.com/white-paper/14556/en/> >
- [4] TU Ilmenau Service GmbH, *Hardware Manual - m:explore. M-sequence ultra wideband evaluation kit*. Ilmsens, Germany, 2016.
- [5] Herrmann, R., *HAL-API for using ILMsens m:explore UWB sensors with TE0630 USB backend*. API documentation, Ilmsens, Germany, 2016.
- [6] Rovňáková, J., *Complete signal processing for through wall target tracking by M-sequence UWB radar system*. Dizertačná práca, Technická univerzita v Košiciach, Košice, Slovensko, 2009.
- [7] Application Design Patterns: State Machines. [online] [citované 2021]. Dostupný na internete: < <https://www.ni.com/cs-cz/support/documentation/supplemental/16/simple-state-machine-template-documentation.html> >
- [8] Event-Driven Programming in LabVIEW. [online] [citované 2021]. Dostupný na internete: < <https://www.ni.com/cs-cz/support/documentation/supplemental/06/event-driven-programming-in-labview.html> >

# Bezpečnostné riziko pri generovaní náhodných dát v prostredí virtuálnych strojov s OS Windows

<sup>1</sup>Marek ROHAČ, <sup>2</sup>Miloš DRUTAROVSKÝ

<sup>1,2</sup> Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>marek.rohac@student.tuke.sk, <sup>2</sup>milos.drutarovsky@tuke.sk

**Abstrakt** – Cieľom článku je poukázať na bezpečnostné riziko, ku ktorému dochádza počas generovania náhodných čísel tesne po obnovení snímky obrazu virtuálneho stroja s operačným systémom Windows. Uvedené procesy sú demonštrované pomocou opisu a experimentu s aktuálnymi nástrojmi.

**Kľúčové slová** – RNG, Virtuálne stroje, Windows, winapi, OpenSSL

## I. ÚVOD

Virtuálny stroj (ďalej VM) je softvér, pomocou ktorého dokážeme vytvoriť abstraktné, respektíve virtuálne prostredie medzi našim počítačom a inou platformou. Celý tento proces je nezávislý od aktuálne použitého operačného systému (ďalej OS). Používateľ dokáže takto spustiť aj aplikačné sady, ktoré nie su určené práve pre jeho OS. VM umožňuje inštaláciu virtuálneho OS na aktuálnom systéme pomocou zdieľania hardvérového vybavenia počítača. V dnešnej dobe je virtualizácia veľmi populárna. VM sa uplatňujú v rôznych sférach. Príkladom použitia sú spoločnosti zamerané na virtualizáciu sieti. Prostredníctvom jednoduchej náhrady alebo rozšírenia softvéru dokážu poskytnúť väčšiu spoľahlivosť a flexibilitu svojich služieb, bez potreby nákupu viacerých zariadení. Obdobne zasahuje virtualizácia aj do procesu vývoja a testovania aplikácií.

Problematika VM je natoľko rozsiahla, že by mohla byť predmetom osobitného rozsiahleho článku. Avšak vzhľadom k aktuálnej limitácii veľkosti obsahu je opis základných princípov a metód týchto nástrojov vynechaný. Pre čitateľa však je k dispozícii odkaz na video [1]. Autor v priebehu úvodných minút opisuje základné princípy virtualizácie. Následne aj demonštruje spustenie konkrétnej linuxovej distribúcie operačného systému. Ak je čitateľ v tejto oblasti nový, mal by zvážiť prezretie videa. Dôvodom je, že pri ďalšom postupe sú použité termíny typické pre platformu VM.

V publikácii [2], autori zrealizovali úspešný útok vo virtuálnom prostredí s OS Windows. Dokument opisuje útok na implementáciu TLS protokolu pri štandardne zabezpečenej komunikácii medzi serverom a klientom. TLS (z ang. Transport Secure Layer) sa používa na zabezpečenie komunikácie s využitím certifikátov pri distribúcií verejného kľúča. Absolútnym základom útoku je vedomosť tzv. bezpečnostnej zraniteľnosti pri obnovení snímky obrazu (z ang. VM reset vulnerabilities). Následne autori dokázali pri opakovanom obnovení snímky extrahovať tajný kľúč servera. Ten vznikne deterministickým procesom pričom sa použijú náhodné dáta z generátorov náhodných čísel (z ang. Random Number Generator, ďalej RNG). Vlastník tohto údaju sa môže vydávať za server a klient nedokáže rozoznať rozdiel.

Tento článok demonštruje aktuálny bezpečnostný problém v prostredí OS Windows na platforme VM, avšak so zameraním na proces generovania náhodných dát po obnovení snímky obrazu s uvedeným OS. Vzhľadom na využitie výstupov z RNG je teda teoreticky možné použiť uvedený problém aj pri pokuse o sieťový útok. Overenie aktuálnosti problému je realizované experimentom. Použité sú RNG rozhrania systému Windows (BCryptGenRandom) a nástroja OpenSSL (príkaz rand), na generovanie kryptograficky bezpečných pseudo-náhodných čísel. Práve tieto dáta sa používajú ako inicializačná hodnota (z ang. Seed) pre generátory používané v kryptografických algoritmoch.

## II. OPIS BEZPEČNOSTNÉHO RIZIKA

### A. Snímky obrazov v prostredí VM

Pred opisom problému je nutné vysvetliť pojem snímka obrazu. Virtuálne stroje poskytujú možnosť spustiť ľubovoľný OS na počítači, bez nutnosti zmeny aktuálneho systému. Ďalšou z vymožeností tohto prostredia je možnosť vytvorenia kópie aktuálneho stavu pomocou tzv. snímky obrazu (z ang. Snapshots). Pri tomto úkone dochádza k úplnému uloženiu stavu daného systému. Výsledkom je súbor, vďaka ktorému je možné kedykoľvek obnoviť systém do stavu, aký bol počas vytvorenia snímky. Vráťane všetkých údajov v operačnej pamäti. Využitie je výhodné najmä pri spúšťaní alebo inštalácii programov z neznámych zdrojov. Tie môžu poškodiť OS. Ďalším príkladom je aktualizácia systému. Po poškodení OS je takto možné obnoviť systém do bodu kedy bolo všetko v poriadku. Po načítaní snímky je možné ďalej pokračovať v práci, avšak iba s dátami z obdobia vzniku snímky.

### B. Zraniteľnosť systému pri obnovení snímky obrazu

VM reset vulnerabilities pomenúva pôvodný anglický názov pre bezpečnostné riziko. Doslovný preklad do slovenčiny nie je celkom výstižný. Preto zavedieme pomenovanie problému ako – Zraniteľnosť pri obnovení snímky obrazu VM. Po reštarte uvedeným spôsobom, je možné zreprodukovat' aj výstupy systémových CSPRNG, viac informácií o RNG nájde čitateľ v [3, kap. 1]). Táto skutočnosť rapídne znižuje bezpečnosť kryptografických algoritmov ako AES (z ang. Advanced Encryption Standard) [4, kap. 5], ktorý sa používa pri symetrickom šifrovaní, resp. dešifrovaní dát práve pomocou tajného kľúča. Ďalším príkladom je taktiež DSA (z ang. Digital Signature Algorithm) [4, kap. 11], používaný pri digitálnych podpisoch. Ak má útočník k dispozícii kľúče, s ktorými uvedené nástroje pracujú, ich použitie za účelom zabezpečenia komunikácie je bezvýznamné. Pri frekventovanom používaní rovnakej snímky obrazu si teda útočník môže všimnúť podobnosti kľúčov.

## III. EXPERIMENTÁLNE OVERENIE UVEDENEJ ZRANITEĽNOSTI

### A. Špecifikácia použitého hardvéru

Realizáciu experimentu zabezpečí prenosný počítač. Tabuľka 1 znázorňuje jeho špecifikáciu

Komponenty	Konfigurácia
Názov	ASUS TUF A15
Model	F506IU-AL006T
Verzia OS	Win 10 Home; 64-bit.; v.20H2
Zostava OS	19042.964
CPU	AMD Ryzen 7 Mobile 4800H
RAM	16GB DDR4 2x1600MHz
Úložisko	SSD OM8PCP3512F-AB

Tabuľka 1  
Technická špecifikácia počítača

### B. Špecifikácia použitého softvéru

Pri tomto experimente je použitá iná metodika práce ako v [2]. Vykonanie jedného z pokusov zabezpečí CSPRNG API operačného systému Windows 10. Konkrétne funkcia BCryptGenRandom [5]. Výstupy následne odhalia bezpečnostné riziko.

Prvé kroky sú spojené s voľbou hypervízora. Voľne dostupný nástroj je VirtualBox [6]. Stiahnutelný na: <https://www.virtualbox.org/wiki/Downloads>. V čase experimentu je dostupná verzia 6.0.24 s podporou 64-bitových OS. Nasleduje voľba operačného systému. Použitý je konkrétne Windows 10 Pro vo verzii 1909 (zostava 18363.592, 64-bitový). Konfigurácia systému prebieha podľa video návodu [7]. Dodatočne došlo k úprave možnosti zdieľania súborov na obojsmernú a navýšeniu počtu priradených procesorov (z ang. Central Processing Unit, ďalej CPU) na počet 4. Dôvodom je samotný OS, ktorý sa v prípade použitej konfigurácie zdal pomalý na prácu dvoch CPU. Tieto zmeny však neovplyvnia výsledok experimentu.

V tomto návode autor priradil OS v prostredí VM 2 procesory (z ang. Central Processing Unit, ďalej CPU). Pri konfigurácii je tento počet CPU zmenený na 4. Dôvodom je samotný OS, ktorý sa v prípade použitia konfigurácie so 4 CPU zdal rýchlejší a svižnejší. Zároveň došlo aj k úprave možnosti zdieľania súborov na obojsmernú. Uvedené zmeny však neovplyvnia výsledky experimentu.

### Prostredie Guest VM

Po úspešnej inštalácii obrazu a sprístupnení OS je nutná príprava prostredia.

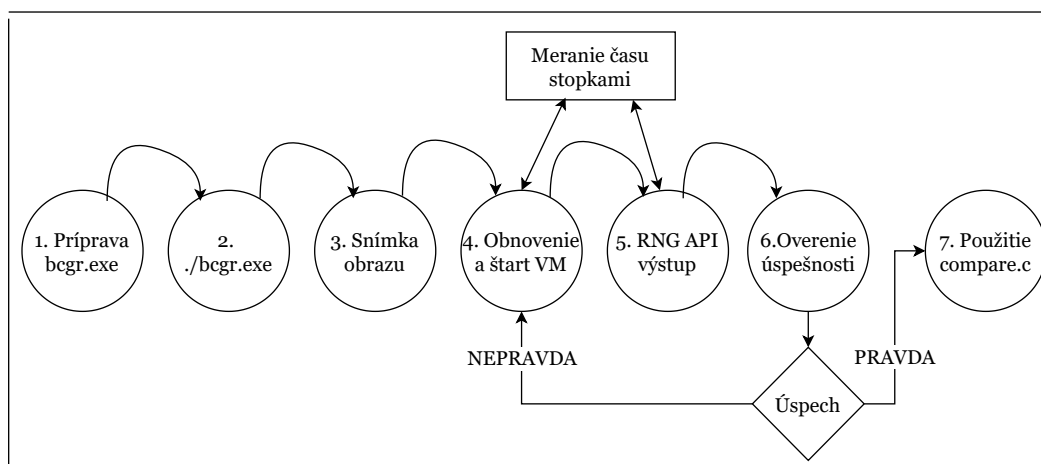
Vykonané inštalácie::

- GCC prekladač – Winlibs GCC, 10.2.0, 64-bitová verzia, pre potreby prekladu programov do strojového kódu. Dostupný na url: <https://winlibs.com/>.
- VirtualBox Extension Pack – 6.0.24, zabezpečí ovládače pre VM,
- Visual Studio Code – prostredie na úpravu, kompiláciu, a aj spustenie programov pomocou príkazového riadka.

### C. Postup pri realizácii experimentu

Pokus je síce realizovaný prostredníctvom odporúčaného rozhrania na generovanie náhodných čísel – BCryptGenRandom [5]. Avšak akékoľvek systémové RNG API dosahuje rovnaké výsledky. Aktuálne má používateľ k dispozícii ešte 2 rozhrania. Konkrétne CryptGenRandom [8] a RtlGenRandom [9]. Všetky uvedené sú opísané a testované v práci [3, kap. 5].

Podstata experimentu spočíva v presnom zopakovaní tých istých krokov od načítania snímky obrazu v približne rovnakom čase. Grafický postup znázorňuje schéma 1. Obsahom nasledujúcej časti je slovný opis. Postup:



Obr. 1 Grafické znázornenie krokov postupu

- 1) Príprava a preklad jednoduchého programu bcgr.exe (BCryptGenRandom). Jeho úlohou je generovanie náhodných čísel pomocou systémového rozhrania Windows. Vid' ukážku zdrojového kódu tohto programu 1. Veľkosť výstupu vygenerovaných dát je v kóde stanovená na 1 920 B. Táto hodnota bola vykonaná pri poslednom experimente. Získaná je postupným násobením dvomi od veľkosti buffer-a 15 bajtov. Teda 15-30-60-120-...-1920. Pri hodnote 1920 B sme s navyšovaním skončili.

Zdrojový kód č. 1 Ukážka zdrojového kódu použitého pri generovaní dát (bcrypt.c)

```
#include<stdio.h>
#include<windows.h>
#include<ntstatus.h>
//gcc bcrypt.c -o bcrypt -lbcrypt
int main() {
    int size= 1920; //size of data to generate in bytes
    BYTE pbData[size];
    BCryptGenRandom(NULL, pbData, size,
        BCRYPT_USE_SYSTEM_PREFERRED_RNG);
    for( int i = 0; i < size; i++ )
        printf("%02X", pbData[i]); //02X for same printing
    printf("\n");
    return 0;
}
```

- 2) Po príprave algoritmu je nutné aby sa používateľ presunul do bodu, tesne pred spustením programu bcgr, v príkazovom riadku. Pri experimente je použitý Windows PowerShell, prostredníctvom editora Visual Studio Code [?], v ktorom sme vykonali prípravu aj preklad programu.

- 3) V uvedenom bode je nutné vytvoriť snímku obrazu. OS bežiaci vo virtuálnom prostredí je možné následne vypnúť.
- 4) V tomto bode používateľ potrebuje pre dosiahnutie úspešnosti používať merací nástroj, napríklad stopky. Meranie je potrebné začať v ľahko identifikovateľnom mieste. Pri pokuse je aktivácia stopiek vykonaná v okamihu spustenia obnovenej snímky OS. Po sprístupnení systému sme spustili program `bcgr.exe`. V čase spustenia programu došlo k zastaveniu merania času. Dôležité pre úspešný experiment je dosiahnuť spustenie `bcgr.exe` vždy v rovnakom čase. Tento krok nie je celkom jednoduchý a vyžaduje zručnosť používateľa. Pri pokuse sme program spúšťali približne 2 sekundy po sprístupnení systému. V čase od začiatku načítavania obrazu až po zapnutie programu nebolo počas experimentu vykonané nič. Žiaden pohyb myši ani stlačenia klávesnice. Toto je aj dôvod, prečo je snímka obrazu zaznamenaná pred zapnutím programu `bcgr.exe`.
- 5) Takto vzniknuté dáta je nutné uložiť mimo VM. Dôvodom je, že pri obnovení pomocou snímky dôjde k zahodeniu všetkých dát, ktoré nie sú jej obsahom. Vráťane dát vzniknutých v kroku 4.
- 6) V prípade neúspešnej realizácií experimentu je potrebné opakovať tento proces od kroku 4. Opakovanie je ukončené v prípade, že daná postupnosť už bola vygenerovaná. Stačí, ak sa používateľ zameria na prvé znaky výstupu. Dôvodom je, že ak by došlo k zmene len jedného bitu pred generovaním, tak výsledná postupnosť bude úplne iná. Tato vlastnosť je charakteristická pre generátor AES, ktorý OS Windows používa. Viac informácií o generátoroch použitých pri procese vytvárania náhodných čísel je dostupných v [3, kap. 2.3].
- 7) Hneď po zaznamenaní zhody v počiatočných hodnotách podrobíme dáta testovaniu zhody. Za účelom tohto procesu je vytvorený program `compare.exe`. Jeho úlohou je kontrola zhody v dvoch rovnako dlhých reťazcoch. Zdrojový kód je znázornený v ukážke 2. Program `compare.exe` je nutné pri každej zmene kompilovať pomocou GCC prekladača. Za zhodu je považovaná rovnosť hodnoty a polohy vo vygenerovanej postupnosti. Výstup poskytuje hodnotu nájdenej zhody v percentách. V prípade úplnej zhody je používateľ upozornený slovným výstupom.

Zdrojový kód č. 2 Ukážka zdrojového kódu `compare.c`

```
#include <stdio.h>
#include <string.h>
//gcc compare.c -o compare
/*
place for generated data
*/
int main(){
    char string1[]={"1.dataToCompare"};
    char string2[]={"2.dataToCompare"};
    int size=sizeof(string1)/sizeof(string1[0]);
    if(strcmp(string1, string2)==0){
        printf("ABSOLUTE MATCH\n");
    }else {
        int i=0,counter=0;
        while(i<size){
            if(string1[i]==string2[i])
                counter++;
            i++;
        }
        printf("MATCH in %.2f percent\n",
            (double) (counter*100)/size);
    }
    return 0;
}
```

### Experiment s knižnicou OpenSSL

Po úspešnom prvom experimente so systémovými rozhraniami je overená aj knižnica OpenSSL. Obsahuje množstvo implementácií kryptografických algoritmov a zároveň poskytuje kompatibilitu pre zariadenie, nezávisle od OS. Pri generovaní náhodných dát je použitý ako zdroj náhodnosti výstup z CSPRNG API systému, na ktorom je naimplementovaná zvolená funkcionálna. Teda ak je niektorá z funkcií na RNG implementovaná v OS Windows, tak vstupom OpenSSL generátora je výstup funkcie `BCryptGenRandom`. Periodicky dochádza k reseed procesu, pričom prvotná inicializácia prebehne pri spustení.

Táto časť používa nástroj OpenSSL [10], vo verzii 1.1.1k. Pri experimente boli aplikované rovnaké metódy ako pri softvérových rozhraniach. Generovanie sme vykonali v rozhraní OpenSSL, konkrétne cez prostredie Command Prompt (ďalej cmd). Knížnicu OpenSSL sme spustili príkazom `openssl` v prostredí cmd. V spustenej aplikácii sme použili príkaz:

```
rand -out result.txt -base64 100.
```

Pred vykonaním došlo k vytvoreniu snímky obrazu. Úlohou príkazu je vygenerovať náhodné dáta s veľkosťou 100 bajtov a uložiť ich do textového súboru `result.txt` vo formáte Base64 [11]. Pokus sme zopakovali 30-krát avšak ani raz sa nám pri experimente nepodarilo zreprodukovat' výstup z generovania.

#### IV. VYHODNOTENIE VÝSLEDKOV TESTOVANIA

Vzhľadom k dobe, ktorá ubehla od uverejnenia článku [2], bolo očakávané úplné vyriešenie tohto bezpečnostného rizika. Počas experimentu sa podarilo čiastočne vyvrátiť túto hypotézu. Dôkazom je videozáznam, ktorý demonštruje postup od 4. do 7 kroku experimentu. Záznam je obsahom prílohy A v práci [3], spoločne so zdrojovými kódmi použitými pri experimente. Zároveň ich čitateľ dokáže vyhľadať na webovej platforme Youtube – [12]. Vďaka videu je zrejmé, že isté bezpečnostné riziko pretrváva.

Je však nutné poukázať na skutočnosť, že zopakovať výstup generovania pomocou systémových RNG API s rovnakým výsledkom, vyžaduje splnenie podmienok, ktoré užívateľ pri bežnej práci nevykonáva. Vid' vyššie uvedený postup. Obdobne bolo potrebné počas testovania vykonať veľké množstvo pokusov. Niekedy bolo nutné počas generovania konkrétnej postupnosti absolvovať

RNG Funkcia	Pokusy	
	Celkom	Úspešné
<b>BCryptGenRandom</b>	50	29
<b>CryptGenRandom</b>	20	11
<b>RtlGenRandom</b>	20	14
<b>Rand_bytes</b>	50	0
<b>Rand_priv_bytes</b>	50	0

Tabuľka 2  
RNG funkcie a pokusy vykonané pri experimente

reštart obrazu aj viac ako 20-krát. Vid' tabuľku 2.

Uvedený bezpečnostný problém však neplatí v prípade použitia modernej kryptografickej knižnice OpenSSL. Pri testovaní bezpečnostného rizika pri obnovení snímky obrazu v prostredí VM s OS Windows 10 sa ani raz nepodarilo zreprodukovat' výstup z RNG rozhrania v tejto knižnici. Tento fakt potvrdzuje kvalitu spracovania kryptografickej knižnice OpenSSL, ktorá je pravidelne udržiavaná i aktualizovaná. V prípade zistenia nejakej chyby dochádza často k jej rýchlemu odstráneniu.

V porovnaní s [2], je situácia v dnešnej dobe pravdepodobne nenapodobiteľná. Je nutné si uvedomiť, že istý potenciál na útočenie existuje. Ak Windows 10 beží prostredníctvom hypervízora, tzv. Hyper-V, tak po obnovení následne dochádza k obnoveniu vnútorných inicializačných hodnôt (z ang. reseed) stavov hlavného generátora systému. Týmto sa značne redukuje čas, kedy je možné uvedený bezpečnostný problém využiť. Jeho implementácia je dnes štandardná. Okrem toho väčšina softvérových hypervízorov pracuje aj v kooperácii s hardvérovými ako Intel VM-x a AMD-V. Jedinou podmienkou je ich povolenie v BIOS-e. Na dnešných počítačoch je však možnosť virtuálizácie hardvéru štandardne zapnutá. Používateľ teda nie je nútený zasahovať do týchto nastavení.

#### Aktualizácia softvérového vybavenia

V čase tvorby článku došlo k aktualizáciám softvérov použitých v tejto kapitole. Zmeny sa týkajú:

- 1) aktualizácia OS Windows – 10, Pro, 64-bitová verzia, v20H2, 19042.985,
- 2) GCC prekladač – Winlibs, 64-bitový, v11.1.0,
- 3) VirtualBox – 6.1.22 (vrátane doplnkov).

Testovanie bolo vykonané aj pomocou aktualizovaných nástrojov. Bezpečnostné riziko aj napriek uvedeným zmenám stále pretrváva. Aktualizovaný videozáznam ([13]) spoločne s údajmi o nástrojoch je taktiež obsahom prílohy A dokumentu [3].



## V. ZÁVER

Obsah tohto článku bol prevzatý z bakalárskej práce – [3, kap. 6]. Je nutné informovať čitateľa o možnom probléme pri realizácii vyššie uvedených skutočností. Ak je na natívnom OS Windows použitý Windows Subsystem pre Linux vo verzii 2 (WSL2), tak je vysoko pravdepodobné, že nedosiahnete výsledky uvedené v tejto publikácii. Pred aplikovaním metód je potrebné deaktivovať túto funkciu. Dôvodom je interferencia medzi Hyper-V a týmto systémom. Občasne môže spôsobovať aj pády, respektíve nespoľahlivý výkon systému.

Článok demonštruje aktuálny bezpečnostný problém v prostredí OS Windows na platforme VM. Teoreticky je teda možné použiť uvedenú chybu aj pri pokuse o sieťový útok. Tento problém je dlhodobo známy fakt. Vzhľadom k veľkej popularite a rozvoju VM preto očakávame v blízkej dobe úplne odstránenie uvedeného bezpečnostného rizika.

## POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-18-0373 a Vedeckou grantovou agentúrou VEGA v rámci projektu č. 1/0584/20.

## LITERATÚRA

- [1] NetworkChuck, “You need to learn Virtual Machines RIGHT NOW!! (Kali Linux VM, Ubuntu, Windows),” January 2021, [Citované: 23.5.2021]. [Online]. Available: [https://www.youtube.com/watch?v=wX75Z-4MEoM&ab\\_channel=NetworkChuck](https://www.youtube.com/watch?v=wX75Z-4MEoM&ab_channel=NetworkChuck)
- [2] T. Ristenpart and S. Yilek, “When good randomness goes bad: Virtual machine reset vulnerabilities and hedging deployed cryptography,” 2010, [Citované: 23.5.2021]. [Online]. Available: <http://pages.cs.wisc.edu/~rist/papers/sslhedge.pdf>
- [3] M. ROHÁČ, *Generovanie náhodných čísel na platforme Windows*. Bakalárska práca, Technická univerzita v Košiciach, 2021.
- [4] D. Levický, *Kryptografia v informačnej bezpečnosti*. Elfa, 2005.
- [5] “Bcryptgenrandom function (bcrypt.h),” 2018, [Online; Citované: 15.5.2021]. [Online]. Available: <https://docs.microsoft.com/en-us/windows/win32/api/wincrypt/nf-wincrypt-cryptgenrandom>
- [6] “Oracle vm virtual box,” 21.5.2021. [Online]. Available: <https://en.wikipedia.org/wiki/VirtualBox>
- [7] MikeTheTech, “Installing windows 10 on virtualbox 6.1.12 – full process, 2020 – video tutorial,” 2020, [Citované: 17.5.2021]. [Online]. Available: [https://www.youtube.com/watch?v=gKQvaPejxpc&ab\\_channel=MikeTheTech](https://www.youtube.com/watch?v=gKQvaPejxpc&ab_channel=MikeTheTech)
- [8] “Cryptgenrandom function (wincrypt.h),” 2018, [Online; Citované: 15.5.2021]. [Online]. Available: <https://docs.microsoft.com/en-us/windows/win32/api/wincrypt/nf-wincrypt-cryptgenrandom>
- [9] “Rtlgenrandom function (ntsecapi.h),” 2018, [Online; Citované: 12.5.2021]. [Online]. Available: <https://docs.microsoft.com/en-us/windows/win32/api/ntsecapi/nf-ntsecapi-rtlgenrandom>
- [10] “Rand - the openssl random generator,” [Online; Citované: 15.5.2021]. [Online]. Available: <https://www.openssl.org/docs/man1.1.1/man7/RAND.html>
- [11] “Base64,” [Online; Citované: 21.5.2021]. [Online]. Available: <https://en.wikipedia.org/wiki/Base64>
- [12] M. ROHÁČ, “Proof for security problem in windows 10 v.1909 vm environment,” 2021, [Citované: 23.5.2021]. [Online]. Available: <https://youtu.be/NTmtWbgm0gw>
- [13] —, “Proof for security problem in windows 10 vm environment – update for version 20h2,” 2021, [Citované: 23.5.2021]. [Online]. Available: <https://youtu.be/eaPwpNJcQr0>

# NTT s konštantnou dobou výpočtu v post-quantovej kryptografii

<sup>1</sup>Eva KUPCOVÁ, <sup>2</sup>Miloš DRUTAROVSKÝ

<sup>1,2</sup>Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>eva.kupcova@student.tuke.sk, <sup>2</sup>milos.drutarovsky@tuke.sk

**Abstrakt** – Článok opisuje princíp a výhody využitia číselno-teoretickej transformácie (NTT) v post-quantovej kryptografii. Zaoberá sa NTT a jej využitím v post-quantových algoritmoch. Pozornosť je venovaná existujúcej implementácii NTT a jej modifikáciám s cieľom analyzovania efektívnosti jednotlivých modifikácií. Samostatne sa venujeme implementácii NTT, ktorej cieľom je dosiahnuť konštantnú dobu výpočtu transformácie. Uvedené sú výsledky a analýza vykonaných meraní na upravených implementáciách.

**Kľúčové slová** – NTT transformácia, post-quantová kryptografia, Montgomeryho redukcia

## I. Úvod

Kryptografia je oblasť matematiky a informatiky, ktorá má za cieľ utajiť informačný obsah tajnej správy pomocou metód šifrovania, ale netají existenciu tejto správy [1]. Moderná kryptografia sa dynamicky rozvíja s rozvojom výpočtovej techniky. Spája sa s elektronickou formou komunikácie a odosielania informácií. Oblasť kryptografie, ktorá vznikla na základe rozvoja fyziky a kvantovej mechaniky sa nazýva kvantová kryptografia. Kým bezpečnosť klasickej kryptografie sa zväčša opiera o matematický problém, ktorý je pre súčasné technické vybavenie neriešiteľný, tak bezpečnosť kvantovej kryptografie sa môže opierať o nejaký fyzikálny jav z kvantovej fyziky alebo mechaniky.

V posledných rokoch sa výskum venuje vývoju kvantových počítačov, ktoré využívajú kvantové mechanické javy na riešenie matematických problémov, ktoré sú aktuálne pre konvenčné počítače veľmi ťažké alebo neriešiteľné. Ak sa kvantové počítače skonštruujú vo veľkom rozsahu, ohrozí to bezpečnosť mnohých bežne používaných kryptografických algoritmov. Kvantové počítače by dokázali v krátkej dobe prelomiť mnoho kryptosystémov, ktoré sú založené na verejnom kľúči vrátane RSA (Rivest-Shamir-Adleman Algorithm), DSA (Digital Signature Algorithm) a ECC (Elliptic Curve Cryptography). Tieto kryptosystémy sa v súčasnosti používajú na implementáciu digitálnych podpisov a výmenu verejných kľúčov. Zohrávajú rozhodujúcu úlohu pri zabezpečovaní dôvernosti a autenticity komunikácie na internete a v iných sieťach.

Vzniknutý problém zraniteľnosti kryptografických systémov s rozvojom kvantových počítačov rieši post-quantová kryptografia. Zaoberá sa kryptografickými algoritmami, ktoré sú odolné voči útokom pomocou (v budúcnosti vytvorených) kvantových počítačov a sú implementované v existujúcich technických prostriedkoch a protokoloch.

Výberom vhodných post-quantových algoritmov sa zaoberá Národný inštitút pre štandardy a technológie, NIST [2]. Proces výberu a štandardizácia vhodných algoritmov trvá aj niekoľko rokov, a ďalších niekoľko rokov trvá, kým sa môže algoritmus považovať za bezpečný a spoľahlivý. Spoľahlivosť si algoritmus buduje tým, že prechádza procesom kryptoanalýzy a je preskúmaný viacerými odborníkmi v danom obore, a taktiež musí byť schopný odolávať rôznym navrhovaným útokom. NIST inicioval v roku 2016 proces vývoja a štandardizácie jedného alebo viacerých ďalších kryptografických algoritmov pre tvorbu verejného kľúča a pre digitálne podpisy. Cieľom je, aby nové kryptografické algoritmy boli schopné chrániť citlivé vládne informácie v dohľadnej budúcnosti, a to aj po nástupe kvantových počítačov. Proces hodnotenia post-quantových kryptosystémov je podstatne zložitejší

ako hodnotenie kandidátov SHA-3 a AES. Jedným z dôvodov je, že požiadavky na šifrovanie pomocou verejného kľúča a pre digitálne podpisy sú komplikovanejšie. Ďalším dôvodom je, že súčasné vedecké chápanie sily kvantových počítačov zďaleka nie je úplné.

Niektoré implementácie post-quantových algoritmov využívajú pre rýchle násobenie polynómov číselno-teoretickú transformáciu (NTT, z angl. Number Theoretic Transform). Medzi post-quantové algoritmy využívajúce NTT patrí napríklad *NewHope*, *Kyber*, *Titanium* a iné [3]. Pre spomenuté post-quantové algoritmy využívajúce NTT boli vykonané merania pre veľkosti vstupov (polynómov) 512 a 1 024 za účelom určenia ich efektívnosti. Implementácie niektorých NTT sú založené na algoritme Cooley-Tukey [4] a základom INTT môže byť Gentleman-Sande algoritmus [4]. Optimalizáciou NTT je možné dosiahnuť celkové vylepšenie daného post-quantového algoritmu. V niektorých implementáciách NTT sa na optimalizovanie využíva antiperiodická (negacyklická) konvolúcia [5]. Článok opisuje tri konkrétne implementácie NTT a ich dosiahnuté výsledky pre rovnaký počet vstupov 512 a 1 024. Cieľom vytvorených implementácií je dosiahnutie konštantnej doby výpočtu NTT z pohľadu CPU cyklov potrebných na jej vykonanie.

## II. ČÍSELNO-TEORETICKÁ TRANSFORMÁCIA

NTT je zovšeobecnením diskretnej Fourierovej transformácie (DFT). Štruktúra NTT je veľmi podobná DFT, ale je definovaná v konečnom poli  $\mathbb{GF}(q)$ , kde  $q$  je prvočíslo. NTT využíva modulárnu aritmetiku. Základným motívom využitia modulárnej aritmetiky je zabezpečenie tzv. uzavretosti pri operáciách medzi prvkami  $a, b \in \mathbb{GF}(q)$  [6]. NTT umožňuje vykonávať rýchle konvolúcie na celočíselných sekvenciách bez zaokrúhľovacích chýb a tiež je využívaná pri násobení veľkého počtu polynómov alebo dlhých polynómov. Má široké použitie aj v počítačovej aritmetike a v implementáciách niektorých kryptografických algoritmov. Za hlavnú výhodu NTT možno považovať, že po použití efektívneho algoritmu dokáže redukovať výpočtovú zložitosť z  $O(N^2)$  na  $O(N \log N)$ .

Pre použitie NTT je potrebné správne zvoliť modul  $q = kN + 1$ , kde potom veľkosť multiplikatívnej grupy  $G_\varphi$  je  $\varphi(q) = q - 1 = kN$ ,  $N$  je mocnicou 2, a určiť generátor (primitívny koreň) tejto grupy  $g$ . Generátor  $g$  je číslo, pre ktoré platí, že  $g^m \equiv 1 \pmod{q}$  a zároveň  $m = \varphi(q)$  [1].

Pomocou generátora  $g$  je možné vypočítať  $N$ -tý primitívny koreň  $\omega$ :

$$\omega = g^k \pmod{q} \text{ a } \omega^N = g^{kN} = g^{\varphi(q)} = 1 \pmod{q}. \quad (1)$$

Pre daný primitívny  $N$ -tý koreň  $\omega$  v konečnom poli, je NTT definovaná:

$$X[i] = \sum_{j=0}^{N-1} x[j] \omega^{ij} \pmod{q}, \text{ kde } i = 0, 1, \dots, N-1. \quad (2)$$

NTT je opísaná transformačnou maticou  $A$ , ktorej hodnoty v  $i$ -tom riadku určujú koeficienty, ktorými je vynásobený vstupný vektor  $\bar{x} = [x[0], x[1], \dots, x[N-1]]$  pri výpočte  $i$ -teho výstupu  $X[i]$  NTT:

$$A = \begin{bmatrix} \omega^0 & \omega^0 & \omega^0 & \dots & \omega^0 \\ \omega^0 & \omega^1 & \omega^2 & \dots & \omega^{N-1} \\ \omega^0 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega^0 & \omega^{N-1} & \omega^{2(N-1)} & \dots & \omega^{(N-1)^2} \end{bmatrix}. \quad (3)$$

Inverzná NTT (INTT) je definovaná:

$$x[i] = N^{-1} \sum_{j=0}^{N-1} X[j] \omega^{-ij} \pmod{q}, \text{ kde } i = 0, 1, \dots, N-1 \quad (4)$$

a opísať ju možno pomocou inverznej transformačnej matice  $B = A^{-1}$ , ktorej hodnoty v  $i$ -tom riadku určujú koeficienty, ktorými je vynásobený vstupný vektor  $\bar{X}$  pri výpočte  $i$ -teho výstupu  $x[i]$  INTT:

$$B = \begin{bmatrix} \omega^{-0} & \omega^{-0} & \omega^{-0} & \dots & \omega^{-0} \\ \omega^{-0} & \omega^{-1} & \omega^{-2} & \dots & \omega^{-(N-1)} \\ \omega^{-0} & \omega^{-2} & \omega^{-4} & \dots & \omega^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega^{-0} & \omega^{-(N-1)} & \omega^{-2(N-1)} & \dots & \omega^{-(N-1)^2} \end{bmatrix}. \quad (5)$$

### III. IMPLEMENTÁCIA NTT TRANSFORMÁCIE A JEJ OPTIMALIZÁCIE

Na experimentálne účely bola využitá voľne dostupná existujúca implementácia NTT a INTT od Michaela Scotta v jazyku C. Autor vo svojom článku [4] opisuje, niekoľko optimalizácií, ktoré boli využité v danej implementácii. Cieľom jednej z použitých optimalizácií je dosiahnutie konštantného času výpočtu NTT a INTT pre rôzne vstupné dáta. Využité boli existujúce známe optimalizácie a v konečnej verzii bola použitá autorova vlastná optimalizácia výpočtu. Náš experiment spočíval v postupnom odstraňovaní jednotlivých optimalizácií, čoho dôsledkom vznikli tri rôzne implementácie NTT [9]. Vzniknuté implementácie sú označované ako *Lazy*, *ConstantTime* a *Naive*.

Scottova implementácia NTT je založená na algoritme Cooley-Tukey [4] a INTT je založená na Gentleman-Sande algoritme [4]. Využitie týchto algoritmov možno považovať za prvú úroveň optimalizácie. *Naive* implementácia využíva spomenuté algoritmy bez ďalších optimalizácií a cieľom pri odoberaní optimalizácií je dostať sa na úroveň tejto implementácie. Jadrom jednotlivých optimalizácií je úprava tzv. motýlikového (butterfly) kódu v najnútornejších slučkách spomínaných algoritmov.

#### A. Lazy implementácia NTT transformácie

Finálna implementácia obsahuje optimalizáciu, ktorá sa nazýva *Lazy*. Hlavná myšlienka *Lazy* optimalizácie spočíva v upravení funkcie *redc*, ktorá je zobrazená vo výpise 1 a stanovení konštánt [4]. Stanovené konštanty pre výpočet NTT a INTT pre veľkosť vstupu 512 a 1 024 sú potrebné pre dosiahnutie korektného výpočtu. Upravená je taktiež funkcia *modmul*, ktorá je zobrazená vo výpise 2.

Výpis 1 Funkcia modulárnej redukcie

```
int_t redc(int_dt T) {
    uint_t m=(uint_t)T*N;
    int_t V=((uint_dt)m*q+T)>>WL;
    return V;
}
```

Výpis 2 Funkcia modulárneho násobenia

```
int_t modmul(int_t a, int_t b) {
    return redc((int_dt)a*b);
}
```

Optimalizácie vnútornej slučky motýlika v implementácii kódu Cooley-Tukey NTT algoritmu a v implementácii kódu Gentleman-Sande INTT algoritmu využívajú rozdielny princíp. Z toho dôvodu je potrebné upraviť samostatne NTT a INTT pre dosiahnutie ich optimalizácií. Optimalizovaný kód motýlika *Lazy* pre NTT je zobrazený vo výpise 3 a optimalizovaný kód motýlika pre INTT je zobrazený vo výpise 4.

Výpis 3 Lazy optimalizácia pre NTT

```
U=x[j];
V=modmul(x[j+t],S);
x[j]=U+V;
x[j+t]=U+2*q-V;
```

Výpis 4 Lazy optimalizácia pre INTT

```
U=x[j];
V=x[j+t];
x[j]=U+V;
W=U+n*q-V;
x[j+t]=modmul(W,S);
```

#### B. ConstantTime implementácia NTT transformácie

*ConstantTime* implementácia využíva Montgomeryho redukciu [4]. Montgomeryho optimalizácia spočíva v Montgomeryho metóde modulárneho násobenia bez využívania operácie delenia. Na architektúre CPU, kde násobenie je konštantné v čase a zaberá 1-2 cykly CPU, a taktiež je hardvérovo optimalizované, možno skonštatovať, že implementácia NTT a INTT môže byť konštantná v čase, z čoho bol odvodený aj jej názov. Za obmedzenie Montgomeryho metódy možno považovať, že vstupné dáta do NTT musia byť pretransformované do Montgomeryho oblasti a výstupné dáta z INTT musia byť pretransformované späť.

Tak ako pri predchádzajúcej optimalizácii, aj táto sa týka optimalizovaného kódu motýlika pre výpočet NTT a INTT. Implementácia môže byť urýchlená predpočítaním omega tabuliek pre NTT a INTT v Montgomeryho tvare. Upravenú funkciu *redc* je možné vidieť vo výpise 5. Zmena sa netýka funkcie *modmul*. Úprava kódu motýlika pre NTT je zobrazená vo výpise 6 a vo výpise 7 je zobrazená zmena týkajúca sa INTT.

Výpis 5 Funkcia *redc* v implementácii v konštantnom čase

```
int_t redc(int_dt T) {
    uint_t m=(uint_t)T*N;
    int_t V=((uint_dt)m*q+T)>>WL;
    V-=q;
    V+=(V>>(WL-1))&q;
    return V;}
}
```

Výpis 6 Montgomeryho optimalizácia pre NTT

```
U=x[j];  
V=modmul(x[j+t],S);  
W=U+V-q;  
x[j]=W+((W>>(WL-1))&q);  
W=U-V;  
x[j+t]=W+(W>>(WL-1))&q;
```

Výpis 7 Montgomeryho optimalizácia pre INTT

```
U=x[j];  
V=x[j+t];  
W=U+V-q;  
x[j]=W+((W>>(WL-1))&q);  
W=U+q-V;  
x[j+t]=modmul(W,S);
```

### C. Naive implementácia NTT transformácie

Poslednou úpravou je odstránenie Montgomeryho metódy, čím je dosiahnutá tzv. *Naive* implementácia. Základom *Naive* NTT je Cooley-Tukey pseudokód [4] prepísaný do jazyka C. Všetky výskyty *mod q* boli nahradené *%q*, čo v jazyku C predstavuje modulo operáciu (celočíselný zvyšok po delení) [7]. Základom INTT je Gentleman-Sandev algoritmus [4]. V jazyku C je *mod q* nahradené operáciou *%q*. V tejto implementácii, na rozdiel od ostatných, sa nevyužíva funkcia *redc* a funkcia určená na vykonanie modulárneho násobania *modmul* je v základnom tvare. Funkcia *modmul* využívaná v *Naive* implementácii je zobrazená vo výpise 8. Kód motýlika pre výpočet NTT je zobrazený vo výpise 9. Výpis 10 obsahuje kód motýlika INTT implementácie.

Výpis 8 Modulárne násobenie v naive implementácii

```
int_t modmul(int_t a, int_t b) {  
    return (int_t)((int_dt)a*b)%q;}
```

Výpis 9 Naive optimalizácia pre NTT

```
U=x[j];  
V=modmul(x[j+t],S);  
x[j]=(U+V)%q;  
x[j+t]=(U+q-V)%q;
```

Výpis 10 Naive optimalizácia pre INTT

```
U=x[j];  
V=x[j+t];  
x[j]=(U+V)%q;  
W=(U+q-V);  
x[j+t]=modmul(W,S);
```

### D. Úprava implementácie pre 16-bitovú aritmetiku

Merania Scottovej implementácie sme realizovali na rôznych platformách. Okrem iných, merania boli vykonané aj na mikrokontroléri (MCU) STM32F103 s 32-bitovým jadrom ARM Cortex-M3 [8]. Na MCU je potrebné šetriť pamäť, preto sme upravili Scottovu implementáciu tak, aby podporovala 16-bitovú aritmetiku. Počas meraní boli vykonávané NTT a INTT pre veľkosť vstupu 1 024, a keďže pre túto veľkosť vstupu je modulo  $q = 12\,289$ , tak 16-bitová aritmetika je postačujúca, a v MCU sme týmto krokom ušetrili značnú časť dátovej pamäte.

Scottova implementácia definuje dátové typy pre celočíselné hodnoty so znamienkom *int\_t* a *int\_dt*, kde *int\_t* predstavuje základnú veľkosť hodnoty a *int\_dt* predstavuje dvojnásobok danej veľkosti. Základná veľkosť hodnoty môže byť 16, 32 alebo 64 bitov, a v podstate nám hovorí, aká aritmetika je v implementácii použitá. Definované sú taktiež ekvivalentné dátové typy pre celočíselné hodnoty bez znamienka *uint\_t*, resp. *uint\_dt*. Pôvodnú použitú základnú veľkosť 32 bitov sme upravili na veľkosť 16 bitov, a tiež sme upravili dvojnásobok veľkosti. Okrem úpravy dátových typov bolo potrebné prepočítať implementáciu využívané Montgomeryho konštanty pre 16-bitovú aritmetiku. Po úprave dátových typov a prepočítaní konštant sme zistili, že je potrebné ešte prispôsobiť jednotlivé funkcie tak, aby využívali upravené dátové typy. Funkcie boli upravené tak, aby využívali vlastné dátové typy všade tam, kde je to potrebné pre dosiahnutie korektného výpočtu NTT a INTT. Prepínanie medzi 16 a 32-bitovou aritmetikou je realizované dvomi sadami globálnych konštant a je podmienené rekompiláciou.

## IV. EXPERIMENTÁLNE OVERENIE A ANALÝZA EFEKTÍVNOSTI IMPLEMENTÁCIÍ

Pre každú z implementácií bolo vykonané meranie počtu CPU cyklov potrebných na vykonanie NTT a INTT samostatne. Na Intel platforme bola využitá na meranie cyklov RTDSC inštrukcia [9]. Merania vykonali výpočet NTT a INTT 10 000-krát a z nameraných hodnôt bol vypočítaný medián a priemer. Každý výpočet NTT použil ako vstupné hodnoty náhodne generované čísla. Pre získanie čo najadekvátnejších výsledkov boli porovnávané hodnoty mediánu a priemeru jednotlivých implementácií. Využitím mediánovej hodnoty sme sa vyhli ovplyvneniu extrémnymi hodnotami pri porovnávaní. Merania boli samostatne vykonané pre veľkosti vstupov  $N = 512$  a  $N = 1\,024$ .

### A. Dosaiahnuté výsledky implementácií

Merania boli realizované v prostredí OS Linux, Ubuntu 20.04 LTS s CPU Intel Core i7-6700HQ. Namerané hodnoty výpočtov NTT a INTT sa nachádzajú v tabuľke 1. Z nameraných hodnôt možno vidieť, že implementovanie optimalizácií značne urýchlí výpočet NTT a INTT z pohľadu potrebných CPU cyklov procesora na vykonanie jednej transformácie. Pri *Lazy* implementácii je taktiež zaujímavým faktom, že pre veľkosť vstupu 1024 je rozdiel v počte CPU cyklov potrebných na výpočet NTT a INTT minimálny. Pri porovnaní *Naive* implementácie NTT pre veľkosť vstupu 512 s *Lazy* implementáciou pre veľkosť vstupu 512 možno zhodnotiť, že implementované optimalizácie dokázali zrýchliť výpočet o takmer  $\frac{2}{3}$ . Taktiež stojí za povšimnutie vplyv Montgomeryho metódy, ktorá dokáže urýchlíť výpočet NTT o  $\frac{1}{2}$ . Preto možno považovať Montgomeryho metódu za veľmi dobrú optimalizáciu z hľadiska urýchlenia výpočtu, hlavne ak ide o veľké množstvo transformácií. Tento spôsob merania nepotvrdil, že implementácia *ConstantTime* je konštantná na meranom systéme.

Priemer a medián výpočtov NTT a INTT transformácie					
		NTT 512 cykly CPU	NTT 1024 cykly CPU	INTT 512 cykly CPU	INTT 1024 cykly CPU
Lazy	Priemer	12 071	24 555	14 884	25 440
	Medián	10 496	22 268	12 071	22 982
ConstantTime	Priemer	16 779	34 308	15 714	31 739
	Medián	14 918	30 213	13 800	27 892
Naive	Priemer	30 489	55 652	18 601	34 527
	Medián	27 158	50 269	16 624	31 018

Tabuľka 1

Porovnanie implementácií na základe priemeru a mediánu počtu cyklov CPU pre veľkosti vstupu 512 a 1024 pre počet opakovaní 10 000 na OS Linux.

Namerané hodnoty pre *ConstantTime* implementáciu neboli konštantné na platforme OS Linux a preto boli merania realizované taktiež v kernel móde OS Linux (KML). Cieľom meraní v KML bolo odstrániť odchýlky výsledkov meraní, ktoré mohli byť spôsobené bežiacimi aplikáciami na pozadí počas merania na OS Linux. Výsledky meraní v KML ukázali, že odchýlky boli do značnej miery redukované, no ani tento spôsob merania nepotvrdil, že implementácia *ConstantTime* je z pohľadu potrebných CPU cyklov naozaj konštantná.

### B. Overenie konštantného počtu CPU cyklov výpočtu NTT a INTT na vstavanom systéme

Táto realizácia overenia konštantnej doby výpočtu bola vykonaná v prostredí Keil ARM simulátora uVision, kde boli vyhodnocované počty cyklov v uVision simulátore. Pre potreby merania boli vytvorené upravené verzie Scottovej implementácie NTT a INTT a to pre 32-bitovú a 16-bitovú aritmetiku. Pre jednotlivé aritmetiky boli vykonané merania pre implementácie *Naive* a *ConstantTime*, kvôli lepšiemu porovnaniu nameraných výsledkov. Výpočet NTT a INTT bol vykonaný samostatne pre náhodne generované vstupné hodnoty a samostatne pre vstupné hodnoty rovné 0.

Pomocou 32-bitovej aritmetiky na meranom systéme sa nepodarilo dosiahnuť konštantné časy výpočtov NTT a INTT, no výsledky meraní ukázali niekoľko zaujímavých faktov. Implementácia *ConstantTime* je výrazne rýchlejšia oproti implementácii *Naive*. Je to spôsobené využívaním 64-bitovej operácie delenia implementáciou *Naive*. 64-bitová operácia delenia je realizovaná ako softvérová rutina s nekonštantnou dobou vykonávania. Merania boli vykonávané pre veľkosť vstupu 512. Výsledky meraní je možné vidieť v tabuľke 2.

Využitím 16-bitovej aritmetiky na meranom systéme sme dokázali dosiahnuť konštantnú dobu výpočtu NTT a INTT *ConstantTime* implementácie pre rôzne vstupné hodnoty. Namerané hodnoty sú zobrazené v tabuľke 3.

Prekvapujúce sú aj výsledky meraní *Naive* implementácie. 16-bitová aritmetika používa hardvérovo podporované 32-bitové operácie. Výsledky meraní ukazujú zaujímavý paradox, kde *Naive* implementácia dosiahla lepšie výsledky, ako implementácia *ConstantTime*. Je to spôsobené hardvérovou podporou 32-bitovej operácie delenia, ktorá je v rámci *Naive* implementácie využívaná.

Na základe meraní možno potvrdiť, že implementácia *ConstantTime* dosahuje konštantnú dobu výpočtu pre rôzne vstupy v rámci 16-bitovej aritmetiky na meranom vstavanom systéme.

Meranie NTT a INTT na MCU				
	Náhodne generované vstupy		Vstupy rovné nule	
	NTT cykly CPU	INTT cykly CPU	NTT cykly CPU	INTT cykly CPU
<b>Naive</b>	5 587 815	6 248 705	4 699 422	6 133 638
<b>ConstantTime</b>	327 109	289 973	313 797	287 925

Tabuľka 2

Porovnanie implementácií na základe počtu cyklov CPU pre veľkosť vstupu 512 na MCU s využitím 32-bitovej aritmetiky.

Meranie NTT a INTT na MCU				
	Náhodne generované vstupy		Vstupy rovné nule	
	NTT cykly CPU	INTT cykly CPU	NTT cykly CPU	INTT cykly CPU
<b>Naive</b>	362 479	358 536	329 903	346 616
<b>ConstantTime</b>	558 264	469 205	558 264	469 205

Tabuľka 3

Porovnanie implementácií na základe počtu jednotkových cyklov CPU pre veľkosť vstupu 1024 na MCU s využitím 16-bitovej aritmetiky.

## V. ZÁVER

Článok sa venuje číselno-teoretickej transformácii a jej využitiu v oblasti post-quantovej kryptografie. Pretože NTT a INTT sú využité v niektorých post-quantových algoritmoch, je možné optimalizáciou transformácií vylepšiť samotný post-quantový algoritmus. Bola upravená existujúca implementácia NTT a INTT v programovacom jazyku C, tak, aby bolo možné skúmať vplyv jednotlivých optimalizácií. Vykonané boli experimentálne merania na viacerých platformách, pomocou ktorých bolo možné určiť efektívnosť jednotlivých optimalizácií. Výsledky meraní sú zosumarizované prehľadnou formou a je opísaná analýza výsledkov meraní. Experimenty potvrdzujú, že implementácia *ConstantTime* je vhodná pre vstavané systémy a naozaj dosahuje konštantnú dobu výpočtu NTT a INTT. Ak sú splnené podmienky pre konštantnú dobu výpočtu NTT a INTT, je výrazne redukovaná možnosť nasadenia útokov využívajúcich postranné kanály. V rámci samotnej bakalárskej práce bola implementovaná v programovacom jazyku C aj vlastná knižnica pre podporné funkcie na inicializáciu tabuliek a pre výpočet NTT a INTT [9].

## POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-18-0373.

## LITERATÚRA

- [1] LEVICKÝ, D. 2018. *APLIKOVANÁ KRYPTOGRAFIA od utajenia správ ku kybernetickej bezpečnosti*. Košice : elfa, s.r.o. ISBN: 978-80-8086-265-7.
- [2] NIST: *National Institute of Standards and Technology* [online]. [cit.2021-05-21]. Dostupné z: <https://www.nist.gov>.
- [3] NIST. 2020. *Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process* [online]. [cit.2021-05-21]. Dostupné z: <https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8309.pdf>
- [4] SCOTT, M. 2017. *A note on the implementation of the Number Theoretic Transform* [online]. [cit.2021-05-21]. Dostupné z: <https://eprint.iacr.org/2017/727.pdf>
- [5] LONGA, P., M. NAEHRIG. *Speeding up the Number Theoretic Transform for Faster Ideal Lattice-Based Cryptography* [online]. [cit.2021-05-21]. Dostupné z: <https://eprint.iacr.org/2016/504.pdf>.
- [6] DRUTAROVSKÝ, M. 2017. *Kryptografia pre vstavané procesorové systémy*. Technická univerzita v Košiciach. ISBN: 978-80-553-2805-8.
- [7] Technotip. 2020. *Modulus or Modulo Division In C Programming Language* [online]. Dostupné z: <https://technotip.com/7272/modulus-or-modulo-division-in-c-programming-language/>
- [8] Arm KEIL: *STMicroelectronics STM32F103C8* [online]. Dostupné z: <https://www.keil.com/dd2/stmicroelectronics/stm32f103c8/>
- [9] KUPCOVÁ, E. 2021. *NTT transformácia v post-quantovej kryptografii*. Bakalárska práca. Technická univerzita v Košiciach.

# Vizualizácia geopriestorových dát o hovoroch na tiesňovú linku

Jozef DOCI, František BABIČ, Viera PAVLIŠKOVÁ

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Letná 9, 042 01 Košice, Slovenská republika

jozef.doci@student.tuke.sk, frantisek.babic@tuke.sk, viera.pavliskova@tuke.sk

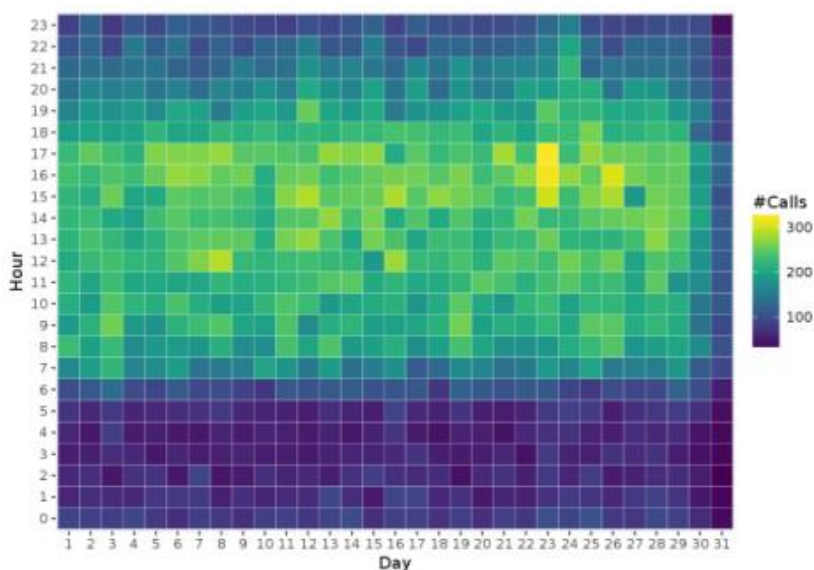
**Abstrakt** — Článok je venovaný analýze a vizualizácii vybranej vzorky geopriestorových dát. Lepšie pochopenie volaní na číslo tiesňovej linky predstavujú dôležitý krok k alokácii a optimalizácii potrebných zdrojov na riešenie príslušných situácií. Získané informácie by mali byť k dispozícii príslušných používateľov (hasiči, policajti, lekárska záchranná služba) v jednoducho pochopiteľnom formáte alebo vizualizácii. V našom prípade sme sa rozhodli vytvoriť prototype takéhoto riešenia vo forme interaktívnej aplikácie využívajúcej voľne dostupné data o volaniach na číslo 911 v oblasti Montgomery County v štáte Pensylvánia, USA. Vytvorená aplikácia ponúka množinu rôznych pohľadov na dáta, ktoré si môže používateľ filtrovať podľa svojich požiadaviek. Jednou z implementovaných funkcií je aj jednoduchá predikcia budúceho incidentu.

**Kľúčové slová** — Geopriestorové data, Vizualizácia, Aplikácia

## I. ÚVOD

Volanie na číslo tiesňovej linky môže byť výsledkom nejakého zranenia, zlého mentálneho stavu alebo ohrozením života či majetku. Nech už sa takéto volanie uskutočňuje z hocikakého dôvodu, je dôležité, aby sa na jeho vyriešenie použili potrebné zdroje. Kuhn a Hoey, ktorí študovali možnosti zásahov záchranných služieb reagujúcich na volania 911 zistili, že ich celkové možnosti boli obmedzené z viacerých dôvodov [1]. Medzi tieto dôvody patrili zastaraný hardvér a softvér, ale aj nedostatok dôveryhodných dát, ktoré by umožňovali zvládať takéto situácie na vysokej úrovni.

Selvam a kol. hľadali podobnosti a vzorce vo volaniach na číslo 911 pomocou K-Means zhlukovania a hotspot analýzy [2]. A tento účel použili dáta o volaniach na číslo 911 z oblasti Montgomery v rokoch 2015 a 2016. Na základe zhlukovania zistili, že najviac hovorov prebiehalo od 08:00 do 18:00. Z pohľadu dní to bolo od 22. do 26. dňa v mesiaci (obr.1).



Obr. 1 Volania na 911 podľa hodiny a dňa v mesiaci [2].

Dôležitým zistením viacerých autorov bola potreba optimalizovať zdroje na riešenie volaní



na základe výsledkov analýzy. Jednou z možností je napríklad zmena v tvorbe stratégií rozmiestnenia policajných, medicínskych a hasičských služieb, a tým zmenšiť reakčný čas na prichádzajúce volania. Druhým dôležitým zistením boli odlišné výsledky na základe rozdielného množstva vstupných dát.

## II. NÁVRH VLASTNÉHO RIEŠENIA

### A. Vizia riešenia

Vizia riešenia spočíva vo vytvorení interaktívnej aplikácie sprostredkujúcej informácie o záznamoch volania na tiesňovú linku, v tomto prípade na číslo 911 v oblasti Montgomery County v štáte Pensylvánia. Základom tejto aplikácie bude spracovanie dát o volaniach, a ich prevedenie z textovej do vizuálnej podoby, napríklad zaznačenie záznamov do máp, interaktívne grafy. Takáto aplikácia nájde využitie v každej zložke záchranných služieb z dôvodu vedenia databázy a informácii o záznamoch ako sú lokácia, adresa, typ, čas a iné odohravajúcich sa v oblasti. Obsahuje funkcie mapovania záznamov a rizikovosti oblastí, na základe ktorých môže predikovať typ a oblasť nadchádzajúcich hovorov. Z tejto predikcie vieme dosiahnuť zníženie reakčného času jednotlivých služieb, znížiť prepravný čas a čas zásahu, ktorý je kritický pri výjazdoch záchranných služieb.

### B. Popis Dátovej množiny

Dátová množina obsahovala 663 522 záznamov a 9 atribútov<sup>1</sup>. Incidentsy pochádzajú z časového intervalu od 10.12.2015 do 29.07.2020. Jednotlivé atribúty sú popísané v nasledovnej tabuľke. Zemepisná šírka a zemepisná dĺžka určujú polohu každého záznamu v množine. Každý záznam, ktorý obsahuje používaná dátová množina sa odohral v meste resp. okolí mesta Philadelphia v štáte Pensylvánia, USA.

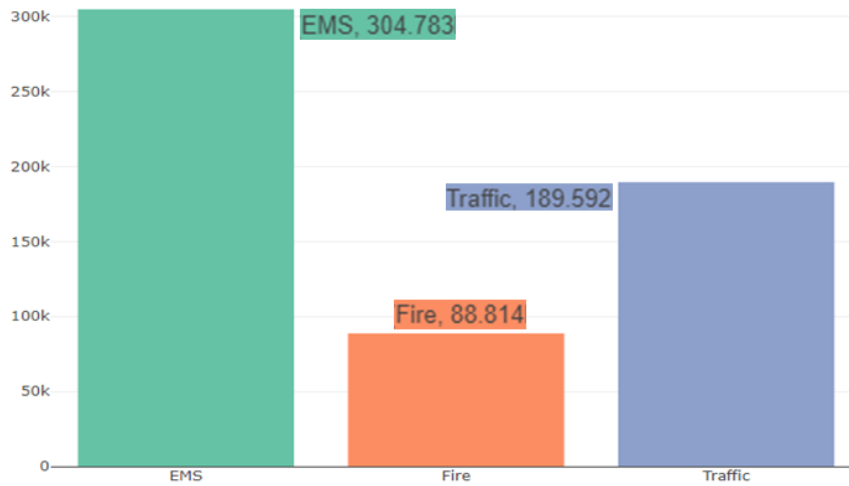
Tab. 1  
 Popis atribútov dátovej množiny

Atribút	Informácie a početnosti					
<b>Popis</b>	Zhromaždené info o adrese, štvrti a čase záznamu					
<b>Typ</b>	<b>EMS</b> – 304 783	<b>Fire</b> – 88 814			<b>Traffic</b> – 189 592	
<b>Názov</b>	Názov incidentu					
<b>Časová stopa</b>	Zhromaždený dátum a čas incidentu					
<b>Dátum</b>	Dátum incidentu					
<b>Mesiac slovne</b>	<b>1.</b> 55 281	<b>2.</b> 48 916	<b>3.</b> 53 239	<b>4.</b> 47 343	<b>5.</b> 50 342	<b>6.</b> 52 576
	<b>7.</b> 52 490	<b>8.</b> 42 299	<b>9.</b> 41 280	<b>10.</b> 44 576	<b>11.</b> 44 078	<b>12.</b> 50 769
<b>Deň v týždni</b>	<b>Pon.</b> 85 258	<b>Ut.</b> 85 888	<b>St.</b> 87 563	<b>Št.</b> 86 779	<b>Pia.</b> 90 327	
	<b>So.</b> 78 148			<b>Ne.</b> 69 234		
<b>Mestská štvrť</b>	Štvrť, kde sa incident odohral					
<b>Adresa</b>	Adresa, kde došlo k incidentu					
<b>Čo sa stalo</b>	Informácia použitá na zaznamenávanie v mapách					

Pre lepšie pochopenie dát sme vytvorili detailné rozdelenie časového záznamu na hodiny, minúty a sekundy. Určili sme zároveň aj deň v týždni a samotný týždeň, ktorý zodpovedal dátumu, kedy sa incident odohral. Distribúcia cieľového atribútu je zobrazená na Obr.2. Môžeme vidieť, že prevládajú incidenty typu EMS (v angličtine “Emergency Medical Services”). Základná množina obsahovala veľa prázdnych záznamov a nepotrebné informácie, ako napríklad ZIP kód. Preto sme odstránili prázdne záznamy a miesta, ktoré mali zle vyznačené súradnice, napríklad boli umiestnené namiesto USA v Indii.

Výsledná množina obsahovala 583 189 záznamov a 19 atribútov.

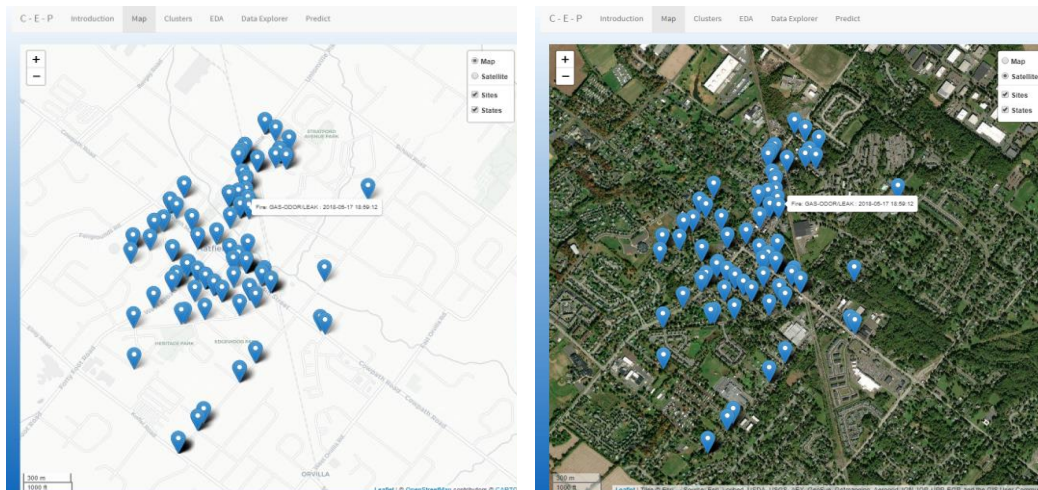
<sup>1</sup> <https://www.kaggle.com/mchirico/montcoalert>



Obr. 2 Distribúcia cieľového atribútu – typ volania.

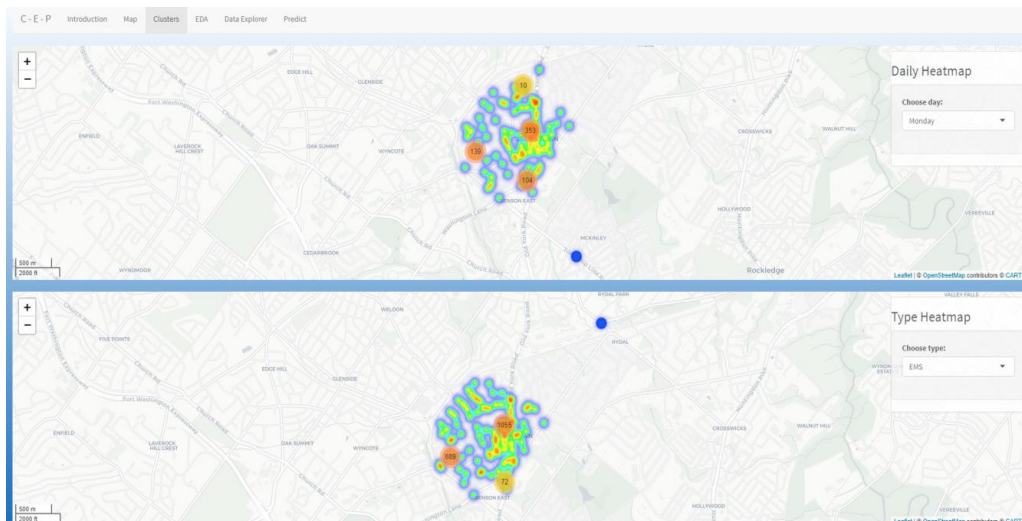
### C. Popis aplikácie

Aplikácia je rozdelená pomocou navigačného panelu na 6 hlavných záložiek: Introduction, Map, Clusters, EDA, Data Explorer a Predict.



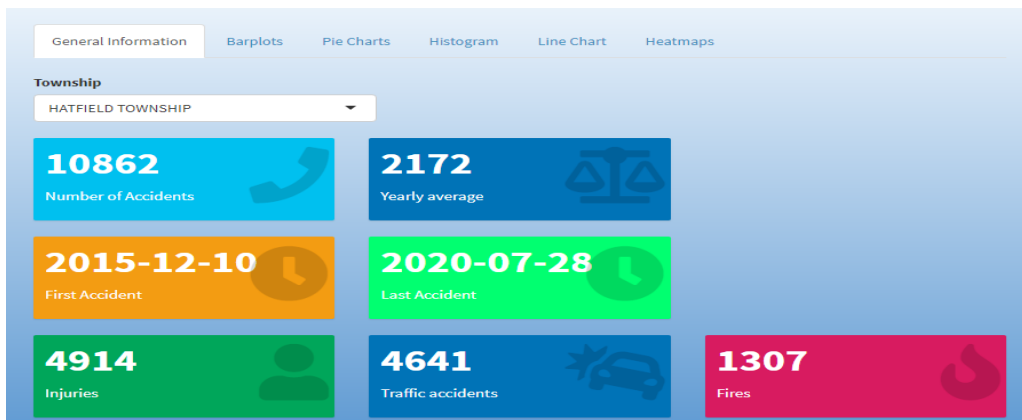
Obr. 3 Záložka Map v aplikácii.

Záložka Map obsahuje mapu a satelitnú snímku oblasti a znázorňuje polohu a popis jednotlivých incidentov, ktoré sa odohrali vo vybranej oblasti.



Obr. 4 Záložka Clusters v aplikácii.

Záložka Clusters okrem navigačného panelu obsahuje heatmapy zobrazujúce zhľuky počtu incidentov. Nachádzajú sa tu 3 mapy, z ktorých každá zobrazuje iný filter, ktorý používateľ vyberá na pravej strane jednotlivej mapy.

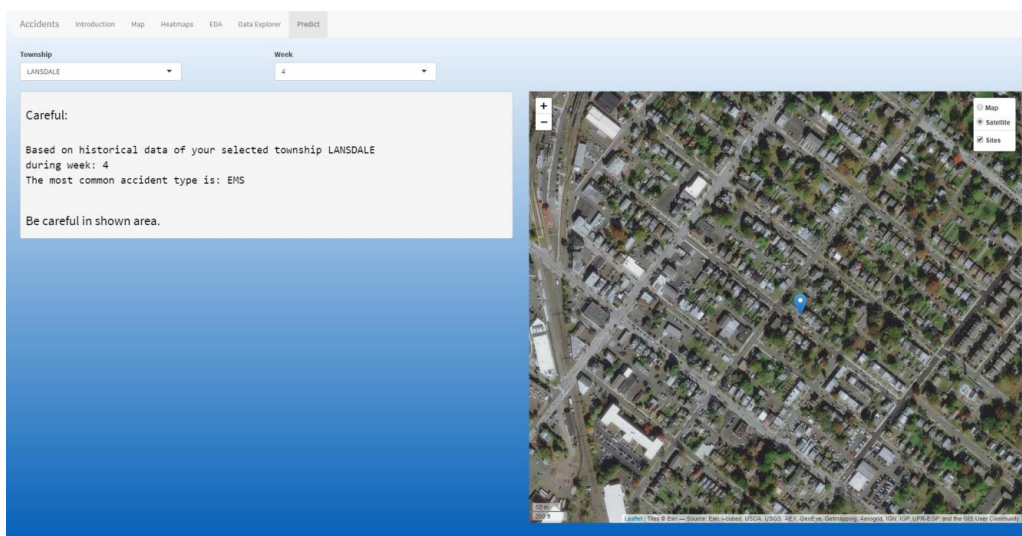


Obr. 5 Záložka EDA (Exploratívna dátová analytika v aplikácii).

Záložka EDA tvorí najrozsiahlejšiu časť aplikácie, ktorá obsahuje šesť nových záložiek:

- General information obsahujúca informatívne boxy podávajúce používateľovi základné informácie o záznamoch vo zvolenej oblasti.
- Záložka Barplots generuje stĺpcové grafy s informáciami, aký počet a v akom pomere sa v danej mestskej štvrti odohrali; rozdelené na množinu incidentov pre danú štvrť za celú dobu, vybraný rok a zvolený mesiac.
- Záložka Pie Charts generuje grafy popisujúce rozdelenie typov hovorov pre používateľom zvolenú štvrť počas celého obdobia dátovej množiny a tiež za vybraný rok. Okrem toho aplikácia zobrazí graf percentuálneho zastúpenia počtu zvoleného typu incidentu počas jednotlivých rokov.
- V záložke Histogram aplikácia poskytne informácie o počte incidentov v jednotlivých mesiacoch za celú dobu dátovej množiny. Používateľ si môže prezrieť celkový počet incidentov v roku rozdelený na jednotlivé hodiny dňa, t.j. ktorá hodina bola najkritickejšia a najčastejšie sa volalo na číslo 9-1-1.
- Záložka Line Charts používateľovi zobrazí vývoj počtu hovorov v zvolenej štvrti počas celého obdobia dátovej množiny a vývoj počtu incidentov počas jedného zvoleného roka.
- Záložka Heatmaps generuje heatmapy zobrazujúce rozloženie počtu hovorov počas jednotlivých dní v mesiaci za celé obdobie, počas zvoleného roka a počet hovorov počas zvoleného mesiaca.

V záložke Data Explorer má používateľ prístup k celej dátovej množine vo forme tabuľky



Obr. 6 Záložka Predict v aplikácii.

Záložka Predict ponúka používateľovi možnosť vybrať mestskú štvrť a týždeň, ktorý ho zaujíma. Následne aplikácia porovná historické záznamy, vyberie najčastejší druh incidentu, ktoré sa v danom časovom intervale vybraného týždňa odohral, na základe triangulácie určí jeho odhadovanú polohu a zobrazí na mape.

### III. ZÁVER

Vytvorená aplikácia ponúka množinu rôznych pohľadov na dáta, ktoré si môže používateľ filtrovať podľa svojich požiadaviek. Jednou z implementovaných funkcií je aj jednoduchá predikcia budúceho incidentu. Dôležitým faktorom pri vývoji aplikácie bola aj výpočtová náročnosť spracovaných dát, ktorú sme riešili pomocou výberu dátových vzoriek podľa skúmaných oblastí.

### POĎAKOVANIE

Táto práca vznikla vďaka podpore Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-16-0213 a Vedeckej grantovej agentúry MŠVVaŠ SR a SAV projekt č. 1/0685/21.

### ZDROJE

- [1] Kuhn, P.; Hoey, T.P., *Improving 911 police operation*. Washington, D.C. National Productivity, pp. 125-130, 1978
- [2] Selvam, Athithyaa et al., *Mining Patterns from 9-1-1 Calls Dataset*. International Journal of Applied Information Systems 11 (2016): 35-40.

# Účinky elektromagnetického poľa na ľudský organizmus

*Jákob ILDŽA, Dušan MEDVEĎ*

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

[jakob.ildza@student.tuke.sk](mailto:jakob.ildza@student.tuke.sk), [dusan.medved@tuke.sk](mailto:dusan.medved@tuke.sk)

**Abstrakt** — Tento článok sa zaoberá problematikou elektromagnetických polí. Popisuje zistené poznatky o tom, ako elektromagnetické pole vplýva na ľudský organizmus, od nízkofrekvenčných po vysokofrekvenčné elektromagnetické polia, a podobne od neionizujúceho po ionizujúce žiarenie.

**Kľúčové slová** — elektromagnetické pole, elektromagnetické žiarenie, nežiadúce účinky, účinky na ľudský organizmus

## I. ÚVOD

Elektromagnetické polia (EMP) vznikajú nielen v okolí elektrických vedení, ale aj v okolí elektrospotrebičov v domácnosti. Postupným technickým rozvojom sa značne zvýšil počet zdrojov elektromagnetického žiarenia. Vznikajú rôzne obavy, či sú EMP nebezpečné, resp. škodlivé. Mnohé výskumy sa zaoberajú vplyvom EMP na ľudský organizmus. Verejnosť je potrebné chrániť pred nepriaznivými účinkami a potencionálnymi zdravotnými rizikami limitovaním expozície EMP. Preto je potrebné detailne poznať zdravotné riziká a ohrozenia a kedy k nim dochádza.

## II. ÚČINKY EMP NA ĽUDSKÝ ORGANIZMUS

Otázky, či je vystavenie EMP prospešné alebo nebezpečné, je stále predmetom mnohých debát a výskumov. Početné štúdie preukázali vplyv EMP na zvieratá (Aydin a Akar 2011, Sonmez 2010), tkanivá a funkčné vlastnosti buniek (Koch 2003, Liburdy 1993), ale zistenia sa stále považujú za predbežné. Naproti tomu veľa štúdií uvádza terapeutické účinky EMP na rôzne orgány a telesné systémy, vrátane zvrátenia kognitívnych porúch Alzheimerovej choroby, stimulácie opravného mechanizmu v kostiach a chrupavkách, hojenie rán a regenerácia nervov. Niektoré z hlavných charakteristík EMP sú jeho frekvencia (teda aj vlnová dĺžka) a intenzita, oba interagujú so živými organizmami rôznymi spôsobmi.

EMP, či už extrémne nízkofrekvenčné (ELF) alebo vysokofrekvenčné (RF), môže viesť k zvýšenému výskytu a riziku rakoviny. Elektromagnetické pole aktivuje bunkové stresové reakcie. Po jej vyvolaní, stresová reakcia vyvolá expresiu génov stresovej reakcie, ako napríklad proteín tepelného šoku (HSP70), ktorý vedie k zvýšeniu hladiny stresových proteínov. Stresové proteíny sa správajú ako chaperóny pomocou opravy a poskladaním poškodených proteínov a ich transportom cez bunecnú membránu. Na vyvolanie tejto stresovej reakcie musí byť aktivované DNA a napriek veľkým energetickým rozdielom medzi ELF a RF sú aktivované rovnaké bunkové dráhy. Konkrétne, DNA sekvencie na promotore (sekvencia DNA, na ktorú sa viaže RNA polymeráza) stresového génu HSP70, reagujú na EMP. EMP taktiež môže priamo pôsobiť na elektróny DNA. Avšak EMP vo vysokofrekvenčnom rozsahu môže viesť k poškodeniu alebo zlomu v reťazci DNA. Akékoľvek rozsiahle poškodenie alebo zmeny DNA, ktoré vyžadujú opravu, môžu viesť k zvýšeniu rizika vývoja rakovinových buniek.

EMP je možné vo všeobecnosti rozdeliť na dva hlavné typy v závislosti od energetickej úrovne elektromagnetických vln. Ak ich elektromagnetické vlny obsahujú dost energie na ionizáciu molekúl, môžu zlomiť väzby medzi molekulami a spôsobujú chemické reakcie. Tieto vlny sú známe ako ionizujúce žiarenie a môžu poškodiť živé bunky, spôsobiť rakovinu, nádory a genetické poškodenia. Vysoké ultrafialové, röntgenové a gama žiarenia sú určité formy ionizujúceho žiarenia. Ak sú úrovne kvantovej energie nedostatočné na to, aby sa zlomili molekulové väzby, tieto elektromagnetické vlny sú známe ako neionizujúce žiarenie. [1] [2] [4]

### III. NEIONIZUJÚCE ELEKTROMAGNETICKÉ ŽIARENIE

Nízka a extrémne nízka frekvencia, rádiofrekvencia a mikrovlny, sú formy neionizujúceho žiarenia. Bežné, umelé zdroje neionizujúceho žiarenia, sú mikrovlnné rúry, počítače, bezdrôtové siete, mobilné telefóny a elektrické vedenia. Všetky tieto vlny vytvárajú istý efekt ohriatia organizmu. Aj malá energia, ktorá je dostupná z najbežnejších zdrojoch, môže spôsobovať akýkoľvek druh poškodenia ľudského tkaniva. Vyššie hustoty EMP, ako napríklad blízko vysokonapäťových vedení a vysieláčov, môžu mať dlhodobé účinky na ľudské zdravie. Hustota elektromagnetických polí nesúvisí iba s výkonom pri zdroji, ale rapídne sa zvyšuje, ak vzdialenosť od zdroja klesá.

Všeobecné obavy v súčasnosti, keďže čoraz viac ľudí používa mobilné telefóny ako nikdy predtým je, že antény mobilných telefónov vyžarujú elektromagnetické žiarenie v blízkosti hlavy človeka. Niektoré štúdie naznačujú tieto potenciálne zdravotné riziká vystavením vysokého výkonu neionizujúceho žiarenia: rakovina, nádor, bolesť hlavy, únava, nespavosť, Alzheimerova choroba, Parkinsonova choroba, impotencia. [1] [2]

### IV. ÚČINKY ELF

ELF (Extremely low frequency) elektromagnetické polia sú generované elektrickou energiou, teda elektrickými strojmi a prístrojmi, prenosovými vežami a vedením vysokého napätia. Frekvencie do približne 300 Hz sú známe ako nízkofrekvenčné a sú súčasťou neionizujúceho žiarenia elektromagnetického spektra. Účinky elektrických aj magnetických polí na biologické systémy sú veľmi kontroverzné.

Pri vystavení ľudského organizmu elektromagnetickým poliam nízkej frekvencie, vytvárajú sa vo vnútri tela elektrické prúdy, ktoré môžu interferovať s vlastnými elektrickými poľami tela, ktoré súvisia s normálnym biologickým fungovaním. Vystavenie nízkofrekvenčnému elektrickému poľu spôsobuje presne definované biologické odozvy, od pocitu mrzutosti po účinky povrchového elektrického náboja. Svalové bunky sú vo všeobecnosti menej citlivé na priame pôsobenie ako nervové tkanivo. Srdcové svalové tkanivo si zaslúži osobitnú pozornosť, pretože zmena vo funkčnosti je potenciálne život-ohrozujúca. Najsilnejší účinok elektromagnetického poľa, pod prahovou hodnotou pre nerv alebo sval, je indukcia magnetických fosfénov. Je to vnímanie slabého mihotavého svetla v periférii zorného poľa na sietnici. Minimálne prahové hodnoty sú okolo 5 mT pri 20 Hz, stúpajúce pri vyšších a nižších frekvenciách. Predpokladá sa, že tento jav je výsledkom interakcie indukovaného elektrického poľa s elektricky excitovateľnými bunkami v sietnici. Takto vzniká výrastok prednej časti mozgu a dá sa považovať za dobrý, ale konzervatívny model procesov, ktoré sa vyskytujú všeobecne v tkanive centrálného nervového systému. Prahové hodnoty pre sily elektrického poľa v sietnici sa odhadujú približne medzi 50 až 100 mV/m pri 20 Hz, stúpajúce pri vyšších a nižších frekvenciách. S týmito hodnotami samozrejme súvisí aj značná neistota.

Niektorí ľudia tvrdia, že sú hypersenzitívny na elektromagnetické polia. Dôkazy zo zaslepených provokačných štúdií však naznačujú, že hlásené príznaky nesúvisia s vystavením elektromagnetických polí. Existujú iba nejednotné a nepresvedčivé dôkazy, že vystavenie nízkofrekvenčným elektrickým a magnetickým poliam spôsobuje depresívne príznaky.

U zvierat možnosť, že vystavenie nízkofrekvenčným poliam môže ovplyvňovať neurobehaviorálne funkcie, bol skúmaný z mnohých hľadísk pomocou viacerých hľadísk expozičných podmienok. Existujú presvedčivé dôkazy, že zvieratá môžu detegovať nízkofrekvenčné elektrické polia, najpravdepodobnejšie v dôsledku účinkov povrchového náboja a môže vyvolať prechodné vzrušenie alebo mierny stres. Vnímanie povrchového elektrického náboja, priama stimulácia nervového a svalového tkaniva a indukcia fosfénov sietnice oka, je dobre definované a môže slúžiť ako základ pre usmernenia EMP.

Okrem toho, existujú aj nepriame vedecké dôkazy, že mozgové funkcie napríklad na vizuálne spracovanie, koordináciu a motoriku, môže byť prechodne ovplyvnené elektrickými poľami. Naproti tomu dôkazy z iného neurobehaviorálneho výskumu dobrovoľníkov vystavených nízkofrekvenčným elektrickým a magnetickým poliam nie sú dostatočne spoľahlivé na to, aby poskytli základ pre expozičné limity pre človeka.

Výsledky rezidenčných, dobrovoľných a zamestnaneckých epidemiologických štúdií naznačujú, že neuroendokrinný systém nie je nepriaznivo ovplyvnený expozíciou 50 až 60 Hz EMP. Týka sa to najmä cirkulujúcej hladiny špecifických hormónov; ako napríklad melatonínu, uvoľňovaného hypofýzou a na množstvo ďalších hormónov podieľajúcich sa na kontrole metabolizmu tela. Laboratórne štúdiá účinkov vystavenia 50 až 60 Hz na nočné hladiny melatonínu u dobrovoľníkov nenašli žiadne efekty. Celkovo tieto údaje nenaznačujú, že nízkofrekvenčné EMP ovplyvňujú neuroendokrinný systém spôsobom, ktorý by vplýval na ľudské zdravie.

Existujú hypotetické úvahy o tom, či expozícia nízkofrekvenčným poliam súvisí s niekoľkými neurodegeneratívnymi ochoreniami. Pre Parkinsonovu chorobu a sklerózu je malý počet štúdií a neexistujú dôkazy o prepojení nízkofrekvenčných polí s týmito chorobami. Výskumu Alzheimerovej choroby a amyotrofickej laterálnej sklerózy je venovaných viac štúdií. Niektoré z nich naznačujú, že ľudia zamestnaní v profesiách pracujúcich s elektrickou energiou, môžu mať zvýšené riziko amyotrofickej laterálnej sklerózy. Doteraz sa nenašiel žiadny biologický mechanizmus, ktorý by vysvetlil spojitosť medzi EMP a týmito ochoreniami. Analýzy podskupín v rámci štúdií posilňujú dojem nekonzistentných údajov.

Značný počet epidemiologických správ uskutočnených najmä v 80. a 90 rokoch naznačujú, že dlhodobá expozícia 50 až 60 Hz magnetickým poliam môže byť spojená s rakovinou. Zatiaľ čo prvé štúdie skúmali rakovinu u detí, neskôr sa výskum zaoberal aj rakovinou dospelých. Pôvodne pozorované asociácie medzi 50 až 60 Hz magnetickými poľami a rôznymi druhmi rakoviny, neboli potvrdené. Pre detskú leukémiu je však situácia iná. Výskum, ktorý nasledoval po prvých štúdiách naznačuje, že môže existovať slabá spojitosť medzi vyššími úrovňami expozície domácností 50 až 60 Hz magnetických polí a rizikom detskej leukémie. Zároveň nebol identifikovaný žiaden biofyzikálny mechanizmus a experimentálne výsledky laboratórnych štúdií na zvieratách a bunkách nepodporujú predstavu, že expozícia takýmto magnetickým poliam zapríčiňuje detskú leukémiu. Všeobecne platí, že štúdie o účinkoch expozície nízkofrekvenčným poliam nepreukázali žiadnu genotoxicitu polí s nižšími hodnotami ako 50 mT. Na rozdiel od epidemiologických dôkazov asociácie medzi detskou leukémiu a dlhodobým vystavením nízkofrekvenčným magnetickým poliam, sú údaje o rakovine zvierat z rozsiahlych celoživotných štúdií takmer všeobecne negatívne. Údaje bunčných štúdií sú všeobecne zhodné so štúdiami na zvieratách.

Existuje niekoľko dobre-zavedených akútnych účinkov vystavenia nízkofrekvenčným EMP na nervový systém. Ako priama stimulácia nervov a svalového tkaniva. Existujú aj nepriame vedecké dôkazy, že mozgové funkcie ako napríklad vizuálne spracovanie, motorika a koordinácia, môže byť prechodne ovplyvnená elektrickým poľom. Všetky tieto efekty majú prahové úrovne, pod ktorými sa neprejavujú a je možné sa im vyhnúť dodržaním príslušných základných obmedzení. Názor ICNIRP je, že v súčasnosti existujúce dôkazy, že dlhodobé vystavenie nízkej frekvencii magnetických polí je spojené so zvýšeným rizikom detskej leukémie, sú príliš slabé na to, aby vytvorili základ pre usmernenia expozície. [1] [2] [4] [7] [8]

## V. ÚČINKY VYSOKOFREKVENČNÝCH EMP

Vysokofrekvenčné a RF signály sa nachádzajú v rozsahu približne od 10 MHz až 300 GHz. RF elektromagnetické polia sú generované prevažne mobilnými telefónmi, inteligentnými zariadeniami, Wi-Fi signálmi, rádiovými stanicami a radarmi. Rádiové a televízne vysielacie sú veľkými zdrojmi RF energie. Antény vysieláčov vyžarujú úroveň výkonu niekoľkých megawattov. Našťastie, umiestnenie týchto antén je na vysokých vežiach alebo budovách, kde sa v ich blízkosti ľudia nepohybujú. Väčšina pevných liniek a bezdrôtových telefónov majú buď viditeľnú alebo integrovanú anténu do tela telefónu. Pretože antény mobilných telefónov sú v blízkosti hlavy, telefóny vytvárajú pri používaní väčšie hodnoty expozície ako iné typy RF systémov. Domové telefóny alebo vysielачky spôsobujú podobné efekty.

Biologické účinky vystavenia RF elektromagnetickým poliam je možné rozdeliť na bunkové, medzibunkové, enzymatické a karcinogénne. Posledné správy WHO, SCENIHR, SSK a BAFU naznačujú, že neexistujú dostatočné dôkazy na preukázanie účinkov vystavenia EMP na zdravie s výnimkou efektu zahrievania tkanív. Niekoľko štúdií ELF a RF bolo vykonaných s genotoxickými, avšak s kontroverznými výsledkami.

Mnoho doterajších štúdií uvádza výsledky týkajúce sa karcinogénnych účinkov RF. Zatiaľ čo niektoré štúdie neurčili akýkoľvek vzťah medzi expozíciou RF a vývojom nádoru, iné hlásili zvýšené rizikové výsledky. Výsledky štúdií sú taktiež nekonzistentné, pokiaľ ide o to, či má expozícia RF karcinogénne účinky. V jednej štúdií, skúmajúcej účinky 3G signálu na ľudské mozgové bunky vystavené frekvencii 1950 MHz po dobu 12, 24 a 48 hodín, bolo zistené zvýšenie pravdepodobnosti výskytu rakovinového nádoru pri vystavení EMP mobilným telefónom presahujúcim 3 hodiny denne. [2] [5] [6]

## VI. ÚČINKY EMP NA DETI

ELF, vyžarované z elektrických vedení vysokého napätia, môže mať vplyv na deti žijúce v ich blízkosti. V skutočnosti môžu byť deti neustále ovplyvňované aj internými domovými rozvodmi na nízkej napät'ovej úrovni. Väčšina výsledkov týkajúcich sa ELF a zdravia detí, je založená na epidemiologických štúdiách detskej leukémie. WHO dospela k záveru, že neexistujú nijaké priame dôkazy, že deti sú zraniteľnejšie voči EMP, pretože veľmi málo štúdií sa zaoberalo touto

témou. Napriek tomu, vzhľadom na neistý vplyv EMP na deti, odporúča WHO všeobecné opatrenia, napríklad minimalizovať vystavenie EMP v školách, škôlkach a všetkých lokalitách, kde sa deti pohybujú podstatné časové obdobie počas dňa.

Za posledných dvadsať rokov vznikli diskusie o tom, či sú deti zraniteľné voči RF. Tieto obavy vznikli na základe technologického pokroku a doby vystavenia súčasných detí mobilným telefónom. Štúdie na ľuďoch a zvieratách priniesli významné výsledky a zistenia týkajúce sa používania mobilných telefónov. Zistenia zahŕňajú účinky na ľudský organizmus ako narušenie spánku, bolesť hlavy, zmeny synaptickej plasticity, cykly nervových buniek.

Kórejské štúdie, realizované v roku 1993-1999 zahŕňajúce 1928 detí s leukémiou a 956 detí s nádorom mozgu zistili, že riziko leukémie bolo 2,15-krát vyššie v skupine žijúcej do 2 km od AM rozhlasových vysielateľov, ako v skupine žijúcej vo vzdialenosti viac ako 20 km od týchto zdrojov. UK Independent exper group on mobile phones vyhlásila, že deti môžu byť zraniteľnejšie voči EMP ako ľudia iného veku. Bolo uvedené, že deti sú vystavené elektromagnetickým vlnám po dlhší čas svojho života ako dospelí a vtedy, pretože je ich nervový systém stále v procese vývoja.

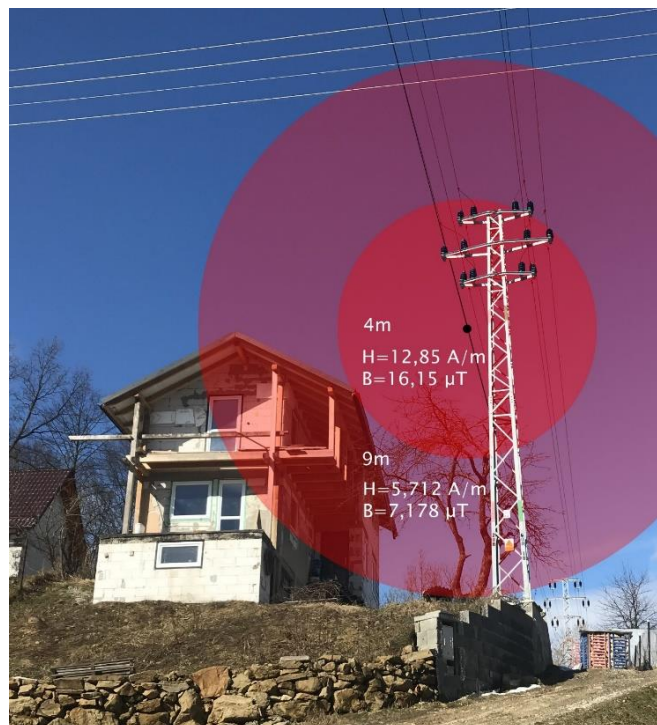
Hrúbka lebky dospelých je približne 2 mm. Avšak hrúbka lebky 5-ročného dieťaťa je približne 0,5 mm a 1 mm v desiatich rokoch. Preto je penetrácia elektromagnetického žiarenia väčšia u detí ako u dospelých. Priemer hlavy dieťaťa je menší, preto energiu absorbujúce oblasti sú citlivejšie na prenikanie RF.

Medzinárodné postupy a poradenstvo odporúčajú tieto postupy na ochranu detí pred EMP: zákaz reklamy alebo predaja mobilných telefónov deťom, uprednostňovanie drôtových dátových sietí v školách a škôlkach pred Wi-Fi, a popísanie špecifickej miery absorpcie.

Zásady predbežnej opatrnosti majú význam vtedy, ak môže ľudská činnosť viesť k morálne neprijateľnej ujme, ktorá je vedecky pravdepodobná. Je potrebné prijať opatrenia, aby sa tomu zabránilo alebo aby sa zmenšili prípadné škody. Pre deti sú potrebné prísnejšie normy až do nadobudnutia vedeckých poznatkov, konkrétne v školských a predškolských zariadeniach, kde sa zdržiavajú dlhšie. [3] [7]

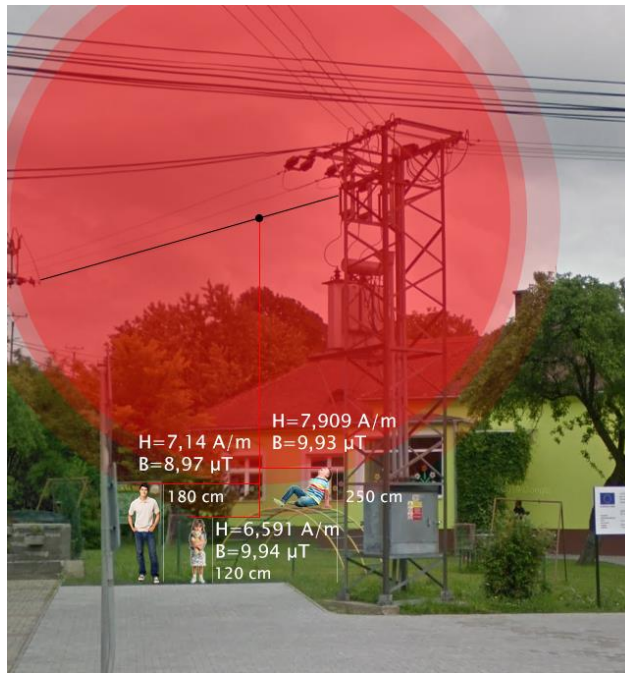
#### VII. PRÍKLADY ROZLOŽENIA EMP V OKOLÍ BEŽNÝCH ELEKTRICKÝCH ZARIADENÍ

Na nasledujúcich grafických zobrazeniach budú uvedené príklady rozloženia EMP v okolí elektrických vedení, ako aj bežne používaných elektrických zariadení v priemysle alebo v domácnostiach. Ďalšie príklady zobrazenia EMP je možné nájsť napríklad v [11].

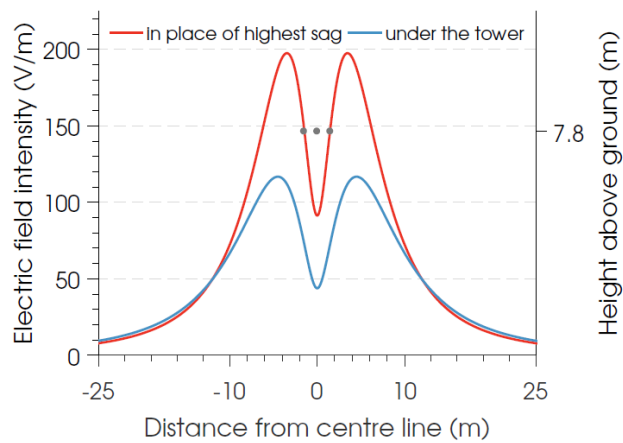


Obr. 1 Rozloženie intenzity magnetického poľa a magnetickej indukcie v okolí rodinného domu

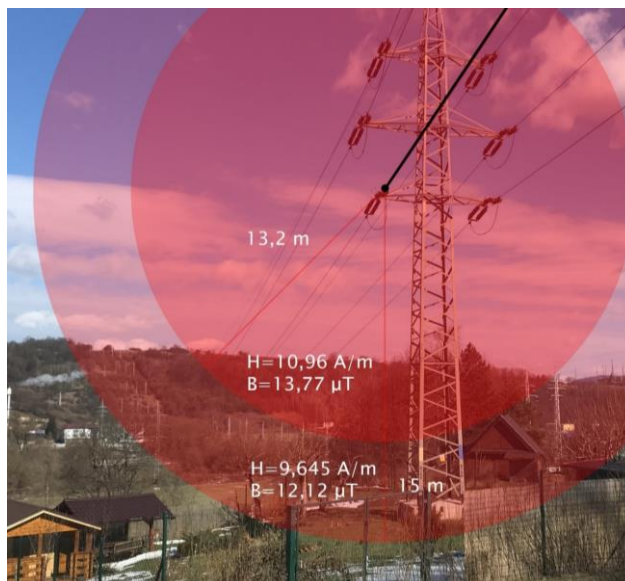




Obr. 2 Rozloženie intenzity magnetického poľa a magnetickej indukcie v areáli základnej školy



Obr. 3 Bočný profil elektrickej intenzity v okolí 22 kV vedenia vo výške hlavy dieťaťa



Obr. 4 Rozloženie intenzity magnetického poľa a magnetickej indukcie na súkromnom pozemku



Obr. 5 Rozloženie intenzity magnetického poľa a magnetickej indukcie v okolí zväracieho zariadenia

### VIII. ZÁVER

Pri ELF EMP boli zistené nežiadúce účinky v podobe blikania svetla v periférii zorného poľa. Výsledky boli nezhodné a nejednoznačné v tom, že nízkofrekvenčné EMP môže spôsobiť depresie. Bolo zistené, že tieto polia ovplyvňujú hladinu hormónov v ľudskom tele. Celkovo tieto údaje nenaznačujú, že nízkofrekvenčné EMP ovplyvňujú neuroendokrinný systém spôsobom, ktorý by vplýval na ľudské zdravie. Obavy, že ELF spôsobuje Alzheimerovu a Parkinsonovu chorobu boli nepotvrdené. Štúdie zaoberajúce sa spôsobením rakoviny elektromagnetickými poľami tieto účinky nepotvrdili. Asociácie medzi 50 až 60 Hz magnetickými poľami a rôznymi druhmi rakoviny neboli potvrdené. Taktiež neboli potvrdené ani dohady, že takéto EMP spôsobujú detskú leukémiu. V prípade RF bolo zistené zvýšenie pravdepodobnosti výskytu rakovinového nádoru pri vystavení EMP mobilným telefónom. Ostatné výsledky zo skúmania RF sú nezhodné na tom, či vysokofrekvenčné EMP spôsobujú rakovinu a nádory. Štúdie zistili nepriaznivý vplyv EMP na reprodukčný systém u mužov. Nízkofrekvenčné EMP môže spôsobiť nepriaznivé účinky na plodnosť a reprodukciu.

### POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu SR a SAV podporenej grantom VEGA 1/0435/19.

### LITERATÚRA

- [1] W ZIMANIAN, A. – HARDIMAN, C.: *Electromagnetic radiation and human health: a review of courses and effects*. July 2005.
- [2] KOCAMAN, A.: *Genotoxic and carcinogenic effects of non-ionizing electromagnetic fields*. 22 February 2018. [online]. [cit.2021-5-15]. Dostupné na internete: <https://doi.org/10.1016/j.envres.2018.01.034>
- [3] MOON, J.: *Health effects of electromagnetic fields on children*. Vol. 63, No. 11, 422-428. 2020.
- [4] MIAH, T. – KAMAT, D.: *Current understanding of the health effects of electromagnetic fields*. Vol. 46, No. 4. 2017.
- [5] BORTKIEWICZ, A. – GADZICKA, E. – SZYMCAK, W.: *Mobile phone use and risk for intracranial tumors and salivary gland tumors – a meta-analysis*. Poland. 2017.
- [6] WUST, P. – KORTÜM, B. – STRAUSS, U.: *Non-thermal effects of radiofrequency electromagnetic fields*. 10 August 2020.
- [7] ZAKI, A. – RAHIM, M. – ZAIDUN, Z.: *Exposure to non-ionizing radiation and childhood cancer: a meta-analysis*. Middle East Journal of Cancer. January 2020.
- [8] ICNIRP: *ICNIRP guidelines for limiting exposure to time-varying electric and magnetic fields (1 Hz – 100 kHz)*. Health physics 99(6):818-836. 2010.
- [9] GÓRSKI, R. – KOTWICKA, M. – SKIBINSKA, I.: *Effects of low-frequency electric field screening on motility of human sperm*. Vol. 27, No. 3. 2020
- [10] SEPEHRIMANESH, M. – DAVIS, D.: *Proteomic impacts of electromagnetic fields on the male reproductive system*. 13 October 2016.
- [11] ILDŽA, J.: *Analýza vplyvu elektromagnetického žiarenia vyžarovaného elektroenergetickými zariadeniami*. Bakalárska práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2021.

# Automation in creation of lists for tests execution in Arena

<sup>1</sup>Miroslav BINAS <sup>2</sup>Patrik KOSCELANSKÝ

Department of computers and informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

<sup>1</sup>miroslav.binas@tuke.sk <sup>2</sup>patrik.koscelansky@student.tuke.sk

**Abstract** – At the Department of Computers and Informatics of Technical University of Košice, an automated evaluation of student programming assignments takes place. *Arena* system carries out the whole process of testing these student submissions. It requires a lengthy configuration file for each problemset, which also needs to be prepared apart from the implementation of tests themselves.

This work describes the creation of a code comments format containing these test configurations, which also aims to unify the tests with all the metadata, in this case, configurations. We also implemented a tool that processes these code comments and generates the *Problemset configuration file*. The solution was evaluated in multiple ways, where most importantly we used the generated configuration in a real-life evaluation process of student submissions of the assignment. Based on this test and the others as well we concluded that this solution is ready to be used in the process of creating the tests of student programming assignments within the given course. We also concluded that the designed format of code comments and the developed tool is easy enough to use by experienced test creators and novices as well.

**Keywords** – Arena, automated evaluation, code generation, configuration, programming assignment

## I. INTRODUCTION

Automated evaluation of submissions of student programming assignments has a long tradition. One of the main reasons for this was the fact that with an increasing number of students attending programming classes, there was also a need to fairly evaluate and assess a big number of student implementations. The first mention of a system that could do something like this comes from the year 1960. In [1] Hollingsworth describes an automatic grader, which could check the output of student programs. From these times a lot of progress was done in this area. The comparison of many tools can be found in reviews such as [2] or [3].

At the Department of Computers and Informatics of Technical University of Košice, such automated evaluation takes place as well in several programming courses. Platform *Arena* was created for this purpose. The system allows fast, safe and objective evaluation of assignments. Staff and volunteers responsible for a specific course, prepare assignments, which students should work on until the deadline and also tests, which are executed on every student submission. Tests are usually implemented as *unit tests*. In order for them to be able to run in *Arena*, the test creators also need to prepare a lengthy configuration file called *Problemset*, which represents something like a "playlist" of what should be executed in *Arena* on every student submission. That means that tests are taking shape in two different places - source files with implementations of tests, but also this configuration file. That brings some amount of duplicity and can result in inconsistencies of tests and their configurations.

Therefore, in this work, we aimed to bring the tests and their configuration into one place. The best approach seemed to be an introduction of some form of *metadata*, which could be placed directly at the place of the implementation of specific *test case*.

## II. PLATFORM ARENA AND CREATION OF TESTS

### A. Structure and functionality of Arena

As was already mentioned, *Arena* consists of several components, which aim to achieve reusability and independence [4]. Specific web services, also mentioned in [5], are the following:

- **Caesar** - Responsible mainly for displaying results of evaluated student submissions. It gets its data from *Gladiator*.

- **Gladiator** - Executes the tests on the student code and evaluates it. Requires configuration file *Problemset*.
- **Conductor** - Executes a structure check of files submitted by students.
- **Spartan** - Web application allowing to solve exercises and assignments in real-time by students.

*Arena* was designed to be universal, which means it supports almost any programming language. It also has two main use cases, which are more closely described in [6]:

- 1) **Scheduled assessment** - In a given time *Arena* downloads student submissions from her *Git* repository. Student's implementation is evaluated by *Gladiator* and results are later available via *Arena* web application.
- 2) **Real-time assessment** - Students solve programming tasks in a web browser environment. Solutions are evaluated on-demand and students can immediately see their results.

To ensure security when executing foreign code, each student's solution is being evaluated in its own *Docker* container. This way, when some fatal error occurs, the whole server isn't affected. More on this issue in [6].

#### B. Creation of tests in order to evaluate student's submissions

In *Arena*, student's solutions are usually evaluated with the help of *unit tests*. Those are carefully prepared by the team responsible for the given programming course. Simplified, the whole process of evaluating a single student submission looks like this:

- At the beginning, the test creator provides his own implementation of the assignment. This "teacher's" solution is considered to be a reference solution, in other words, a correct one.
- Apart from that, he prepares also the implementation of unit tests, against which the student solution will be evaluated.
- These test cases are executed on both solutions with the same input values, where the output from the test creator's program is considered to be the correct one. If the output from the student's solution matches the output from the creator's solution, the student gains points for the given test case.

#### C. Problemset configuration file

In order to be able to run the test cases, *Arena* requires already mentioned playlist, respectively a *Problemset configuration file*, which defines which test cases and how they should be executed in *Docker* container, where student's solution is being evaluated.

*Problemset* consists of individual *tasks*, which are the smallest unit executable by *Gladiator*, e.g. structure check, compilation or unit tests [7]. Usually, one *task* or a *suite* corresponds to one *test case*. *Suite* can contain multiple *tasks*.

### III. USED METHODOLOGY

As we already described earlier, there was a need to automatically create the *Problemset configuration file*, which defines how test cases should be executed in *Arena*. Inspired by the principle *literate programming* described in [8], where combining the implementation with documentation comments is described, we decided to create a format of specially-structured comments, something like *Javadoc*, which would reside at the place of the implementation of tests. To achieve this goal, we identified several standard concepts that we needed to use for a good design of our solution.

#### A. Parsing the configuration comments

Since we wanted to use specially structured comments as a place where our metadata would reside, we identified, that there is a need for grammar for these comments and syntactic analysis.

From a large number of different parser generators we chose library *Lark*<sup>1</sup> for *Python*, which can very easily generate a parser from the provided grammar in *EBNF* format. After parsing it generates *AST*, which can be traversed in order to convert the extracted data into the desired form.

#### B. Validation of parsed metadata

Metadata, which are parsed from the configuration comments need to be checked for allowed values because the parser will only be responsible for checking the format and parsing the data. Values, however, can range. To solve this problem we decided to use a library called *Pydantic*<sup>2</sup>. *Python* classes are created as models, which when initialized, check the allowed data types and ranges of the values. It also provides a possibility to use *validation methods*, which allow a programmer to define more complicated validation rules, or even an implementation to modify the values, for example, to generate a random string.

<sup>1</sup><https://lark-parser.readthedocs.io>

<sup>2</sup><https://pydantic-docs.helpmanual.io>

### C. Code generation

The output of the tool designed in this work should be a *Problemset configuration file*, which is in *YAML* format. For this purposes, there was a need to use some form of *code generation*.

In the book *Domain Specific Languages* [9] code generation is divided into two main categories:

- **Transformer generation** - Based on the input model, the output code is generated.
- **Templated generation** - Variable data are inserted into specially marked places in prepared text.

Final *Problemset* configuration file in order to be usable in *Gladiator*, should have a structure that follows the specification [7]. Therefore, for the purposes of this work, we used *templated generation*. There are many solutions that allow the preparation of templates in order to generate code. We have decided to use *Jinja2*<sup>3</sup>, which is one of the most popular templating system for *Python* programming language used in many tools. Apart from only inserting variable data into a template, it also provides multiple advanced features, such as conditions and inheritance of templates, which can be advantageously used.

## IV. CONCEPT OF A SOLUTION

Regarding the functionality of the tool, we identified a basic flow, which describes how metadata from code comments become the resulting *Problemset configuration file*. The flow is shown in figure 1. It also represents a lifecycle of a single comment in our tool. A closer description of individual steps follows here:

- 1) **Extraction** - All comments are extracted from source files of tests. Comments, which are not relevant for processing are neglected. The text of a comment is saved along with the line number, where the comment starts and the name of a source file, from which the comment was obtained. The output of this step is a list of comments, containing mentioned data.
- 2) **Conversion into dictionary** - In this step, the text form of comments is converted into *Python* dictionaries with the help of the parser. The validity of the data will not be checked at this point. This is a simple check of the grammar and conversion into a different form. All characters from comments which are not needed will be deleted from the text, e.g. ' @ ' characters.
- 3) **Validation** - Correctness of data from comments will be checked against rules defined in *Pydantic* models. Rules are for instance ranges of values or they check for the presence of mandatory values. If any mistake occurs, the user will be warned. The output of this process is a valid object, which is prepared for its conversion into *YAML*.
- 4) **Templated code generation** - In this step, all data are prepared for the generation with a template. Based on the type of tests (including target language), a corresponding combination of templates will be used. Required data are inserted into the template and the resulting text will be written into a *YAML* file.

## V. DESIGN AND IMPLEMENTATION

### A. Problemset configuration file and designed tags

For a good design of the content of the comments, we first needed to understand what the *Problemset* configuration should consist of and study some already existent configuration files. We came to a conclusion that the best form of metadata in comments will be *key - value*, where many keys correspond 1:1 to their meaning in *Problemset* specification [7], but some have slightly different allowed values or are brand new. Since *Javadoc* was our inspiration, we have used similar naming conventions - so we have named keys as the *tags*. In each comment, they start with ' @ ' character and have a value.

*Problemset* consist of both required and non-required keys. Our tool has to be able to generate valid *YAML* which follows the specification, but it doesn't mean, that our designed *tags* and their values have to follow the specification as well. Some values can be generated or derived.

Since *Arena* supports the execution of test in any language, we needed to design our tool in such a way, that it could support these tests as well. We determined that the tool will be extendible to support new kinds of tests. For the start, we mainly focused on supporting the tests that are used in the *Programming* course, which uses *C* language and library *Check* for the tests. That means that for different kinds of tests there can be slightly different required *tags* and their values. Thus, we distinguish:

- 1) **tags universal for any type of testing**,
- 2) **tags specific for a given type of testing** (such as tests in *Programming* course).

All the *tags* that have been designed along with their allowed values are more closely described in *master thesis* [10].

<sup>3</sup><https://jinja.palletsprojects.com>

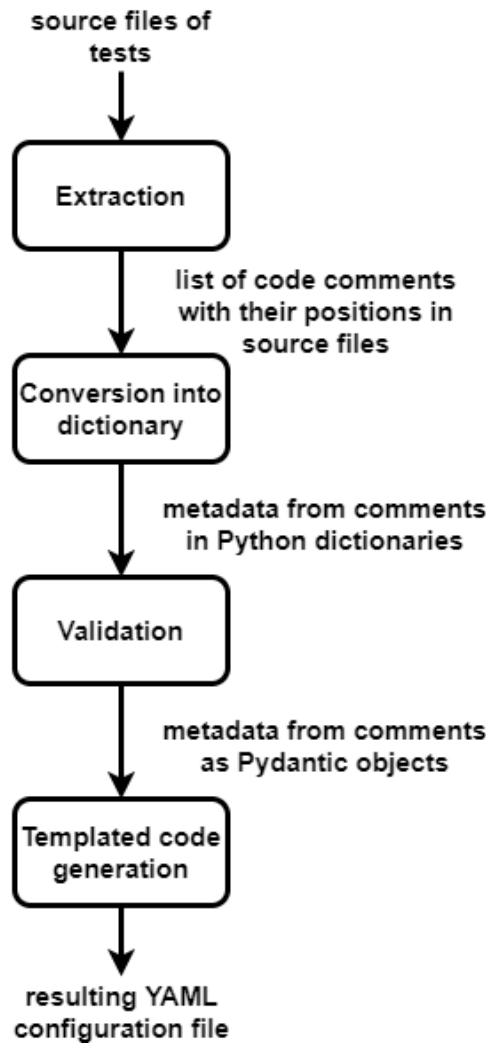


Fig. 1 Data flow diagram of the tool

### B. Comment format and parsing

As we already mentioned, configuration metadata in our specially structured comments are organized as pairs of `@key - value`. The parser won't check if the name of the key or value is valid - that will be checked in the validation step. The `@tag` should be followed by exactly one space followed by `value`. Each comment which is supposed to be processed by the tool should start with a tag starting with `@@type`, where `type` describes what kind of comment it is, e.g. `@@tcase` for a test case. Values can be strings, numbers, boolean values, lists of strings or list of environment variables. An environment variable with its value should be in a form of `ENV_VAR_NAME=value`, where values are always strings, numbers or boolean values. An example of such a comment is shown on the listing 1.

Grammar in *EBNF*-like format is as a string inserted into a constructor of *Lark* parser generator. The generated parser can then process input text and output an *AST*. Parsing is executed for each individual comment. This *AST* is then, thanks to the features of the *Lark* library, traversed and transformed into normal *Python* dictionary. At this point, parsed metadata are prepared for validation.

### C. Validation of extracted metadata

Comment which was extracted and parsed needs to be checked before code generation. The reason is that in order to generate we need to make sure, that before generation starts we have all the values that are required in templates and that extracted values follow the rules and ranges required for a specific type of testing. Resulting *Problemset configuration file* has to make sense for *Gladiator*.

*Pydantic* models are initialized with parsed data, which are checked against several predefined rules. According to these rules, for some values a default value can be used (e.g. `score =`

```
/*
@@tcase

@id "bmp_flip_horizontally_dry_run"
@title "Dry run Test of `flip_horizontally()`"
@description "Test of flipping bmp horizontally."
@hints [
    "Previous height should equal to new width and vice versa."
]
@env [
    CK_RUN_CASE="flip_horizontally(): Dry Run"
]
@strict true
@score 4.5
@features ["valgrind"]
*/
TCase* tc_flip_horizontally_test_flip() { ...
```

Listing 1: Example of structured comments for problemset's test

0.0), or in others cases this value can be generated (e.g. `id` for generated `task`). As was already mentioned, we aimed to design our tool and format of comments in such a way that it could be extendible. For this reason, we provided several *Pydantic* models, but mainly `TestCase` model, which accepts data of all *universal keys*. This model has to be extended for each type of tests (e.g. tests for *Programming* course). In the models which extended this base model, specific rules can be overridden and new validation rules can be added. On the listing 2 the *Pydantic* model for validating universal tags is shown. This model has to be extended.

```
class TestCase(BaseModel):
    id: str = None
    title: str
    description: Optional[str] = None
    score = 0.0
    score_when: Optional[str] = None
    env: dict = {}
    strict: Optional[bool] = None
    hints: Optional[List[str]] = None
    timeout: Optional[float] = None
    replicas: Optional[int] = None
    features: Optional[List[str]] = list()
    raw: Optional[str] = None
```

Listing 2: Pydantic constructor for validation of universal tags

#### D. Code generation

After metadata of configuration comment are validated, the code generation follows. For practical reasons, we decided to leave this implementation as a method of class describing *Pydantic* model. That means that each object containing metadata of exactly one comment will be responsible for generating its own data. Our solution uses multiple *Jinja2* templates, where the most important one generates a `task` or `suite` with `tasks` for exactly one *test case*. For purposes of generating *YAML* code of *test case* we prepared two templates:

- 1) **base test case template**, which contains *YAML* code that doesn't change from one type of tests to another and
- 2) **extending test case templates**, which are prepared for each extension and contain code that is a bit different for each type of tests.

These two templates together generate *YAML* code for a *test case*. Code from extending template is inserted at a specially marked place in the base template and this way it extends it. Extending template usually should contain code for `task`, which describes the execution of a *unit test* and code for additional features, which represent additional execution of the test, usually with some special tools, such as *Valgrind* or *PMD*. If the user specified in a configuration comment, that

such tool should be used (see listing 1, the value of tag `@features`), a task for such tool will be generated as well.

### E. Composing of resulting file

In order for tests to run in *Arena*, it usually isn't enough to provide `tasks` for executing tests. *Problemset configuration file* should have its structure and required data, for each type of testing various "preparation" `tasks` and *YAML* templates should be provided and generated `tasks` should be organized in a manner that makes sense - usually into a `suite` that includes all the tests of an assignment module.

For these purposes, we designed a way to orchestrate the generation and build the resulting *Problemset configuration file*. `ValidationManager` is an abstract class that manages for which tests should `tasks` be validated and then generated. It instantiates *Pydantic* models with parsed data, which validates it and it allows the generation of the final code. Both actions - validation and generation are executed according to a **strategy**, which describes what should be validated and generated. Designed strategies are following:

- **project** - all files with tests should be processed,
- **module** - only files of a given module should be processed,
- **files** - only given files with tests should be processed.

In order for our tool to be able to work with test files and configuration comments in them in the described manner, we needed to restrict how the folder with tests source files should look like (source files of tests organized in subdirectories with the names of the modules) and what the names of these files should look like.

## VI. IMPLEMENTED SOLUTION AND ITS EVALUATION

We implemented our tool as a command-line application, which can be executed directly in the tests directory with different options. By using them the tests creator can control what exactly should be generated. On the listing 3 there is a basic example of using the tool. It defines what type of tests we intend to process (`cprog` as *Programming* and *C*), it defines tests directory and output directory. For other switches, as *strategy*, default options are used. We evaluated this tool along with the created format of comments in multiple ways.

```
$ python -m psgen -mngn cprog -td . -o .
```

Listing 3: Example of use of the tool

### A. Real-life assignment

The created tool was evaluated during *Programming course* on two problemsets. This involves creating structured comments for all the tests, which number was close to 200. From the created configuration comments was then generated *Problemset configuration file* with our tool. Last minor modifications were made to a file and then it was compared with the *Problemset* file that was currently in use. For purposes of comparison *YQ* tool was used. After discrepancies were found, configuration comments were modified and new *Problemset* was generated. This time the resulting file was correct and it could be used in real life. The evaluation of student submissions started and continued to execute without any problems.

### B. User tests

In order to assess the usability of our tool and comments format and also to discover negative aspects of the solution which we didn't recognize during the design and implementation, we decided to also perform user tests. We created a dummy programming assignment, into which the participants were supposed to add configuration comments (provided required metadata). After that, they were supposed to use the tool to generate *Problemset configuration file* in two scenarios. The group consisted of some students, who had previous experience with test creation or configuration, but also from absolute novices. After an introduction into tests creation and an explanation of important aspects of our solution, they were mostly able to finish the tasks. Only a few of them needed a little help.



## VII. CONCLUSION AND FUTURE WORK

We designed a format of specially structured code comments in order to unify tests implementation and their configuration metadata into one place - source files of tests. We also implemented a tool that is able to process configuration comments in tests source files and generate *Problemset* file, which is a playlist of tests which *Arena* executes on student submissions.

After the evaluation we concluded, that our solution is fit to be used for the creation of tests of assignments in *Programming* course. However, our design supports also extending the tool to support testing in different courses. Based on the results of user tests we also concluded that both more experienced test creators and novices as well are able to learn to work with our comments format and the tool quickly enough. After a quick introduction, they were mostly able to solve given tasks without help.

During the evaluation, we also discovered several deficiencies of the solution, which can be better analyzed and improved in future work. Suggestions on how to continue were also provided by user test participants. For example, a creation of an extension for an *IDE*, which would automatically complete or prepare configuration comments.

## REFERENCES

- [1] J. Hollingsworth, "Automatic graders for programming classes," *Commun. ACM*, vol. 3, no. 10, p. 528–529, Oct. 1960. [Online]. Available: <https://doi.org/10.1145/367415.367422>
- [2] K. Ala-Mutka, "A survey of automated assessment approaches for programming assignments," *Computer Science Education*, vol. 15, pp. 83–102, 06 2005.
- [3] P. Ihantola, T. Ahoniemi, V. Karavirta, and O. Seppälä, "Review of recent systems for automatic assessment of programming assignments," 01 2010.
- [4] M. Biñas, "Identifying web services for automatic assessments of programming assignments," in *2014 IEEE 12th IEEE International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 2014, pp. 45–50.
- [5] E. Pietriková, J. Juhár, and J. Štátná, "Towards automated assessment in game-creative programming courses," in *2015 13th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 2015, pp. 1–6.
- [6] J. Stastna, J. Juhár, M. Binas, and M. Tomášek, "Security measures in automated assessment system for programming courses," *Acta Informatica Pragensia*, vol. 4, pp. 226–241, 12 2015.
- [7] KPI TUKE. Problemset. [Online]. Available: <https://arena.kpi.fei.tuke.sk/gladiator-docs/problemset.html>
- [8] D. E. Knuth, "Literate Programming," *The Computer Journal*, vol. 27, no. 2, pp. 97–111, 01 1984. [Online]. Available: <https://doi.org/10.1093/comjnl/27.2.97>
- [9] M. Fowler, *Domain Specific Languages*, 1st ed. Addison-Wesley Professional, 2010.
- [10] P. Koscelanský, "Automatizovaná tvorba zoznamov na spúšťanie testov v systéme arena," Master's thesis, Technical University of Košice, Slovakia, 2021.

# Smerovanie a strojové učenie v MANET

<sup>1</sup>Filip ZVEREC, <sup>2</sup>Ján PAPAĽ

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

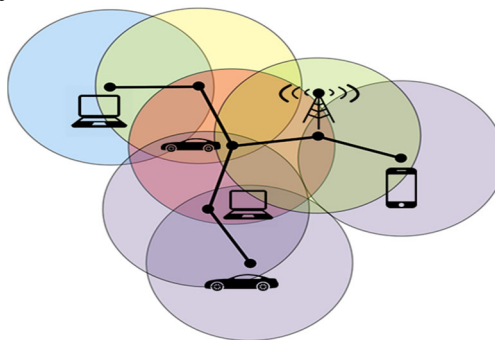
<sup>1</sup>filip.zverec@student.tuke.sk, <sup>2</sup>jan.papaj@tuke.sk

**Abstract** — Tento článok sa zameriava na možnosti implementovania a využívania strojového učenia (ML) v procese smerovania vo viacpreskokových ad-hoc sieťach (MANET). Strojové učenie predstavuje významný míľnik pre samotnú činnosť smerovacích protokolov v MANET. Samotné testovanie algoritmov je realizované na základe analýzy možnosti využitia strojového učenia v procese smerovania v MANET sieťach. V článku sú analyzované 3 typy smerovacích protokolov využívajúce ML v simulačnom prostredí NS-3.

**Keywords** — MANET, strojové učenie, Q- učenie, smerovanie

## I. VYUŽITIE STROJOVÉHO UČENIA V MANET SIEŤACH

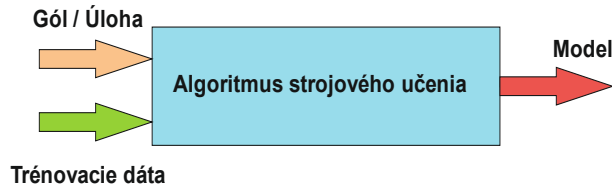
Mobilné ad-hoc siete (MANET) sú bezdrôtové, autonómne samo-organizujúce sa siete bez pevnej topológie. V takejto sieti, slúži uzol v rovnaký čas aj ako vysielateľ aj ako prijímač. Všetky uzly sú si rovnocenné a môžu sa voľne pohybovať, pripojiť či odpojiť zo siete. Mobilné uzly, ktoré sú v rádiovom dosahu s inými uzlami môžu priamo komunikovať a posielat' nevyhnutné informácie. Na takúto komunikáciu majú všetky uzly vstavané bezdrôtové rozhranie. Takýto druh siete je plne distribuovaný a môže fungovať na akomkoľvek mieste bez pomoci akejkoľvek pevnej infraštruktúry akou sú napríklad prístupové body alebo základňové stanice. Rozdiel s WSN sieťami je v tom, že pri MANET sieťach je zmena topológie väčšinou zapríčinená zmenou polohy avšak pri WSN je zmena topológie väčšinou zapríčinená zlyhaním senzora respektíve vybitím batérie a taktiež v sieti neexistuje žiadny zberný uzol, ktorý je väčšinou statický a len funguje ako výstupná brána pre zozbierané údaje. Takéto MANET siete majú najlepšie využitie v oblastiach, kde je zložitý geografický prostredie, ktoré vyžaduje mobilné uzly bez pevnej infraštruktúry. Príkladom môžu byť núdzové situácie, kde záchranné služby nemôžu využiť pôvodnú rádiovú sieť, či už z dôvodu preťaženia siete, jej zničenia, alebo slabého signálu v oblasti. Avšak dynamickosť MANET sietí, frekventované zmeny v sieti zapríčinené mobilitou, otvorené bezdrôtové vysielanie robia tieto siete náchylné voči bezpečnostným útokom. Tento problém vychádzajúci zo základných charakteristík MANET sietí a je neustále testovaný novými bezpečnostnými protokolmi, ktoré sa snažia zachovať výhody týchto sietí, ale nie na úkor ich bezpečnosti. Príklad takejto MANET siete môžeme vidieť na obr. 1 [1].



Obr. 1 Princíp činnosti MANET siete

Vďaka neustále narastajúcemu rozvoju internetu, sieťový výskum pritiahol za posledných niekoľko desaťročí veľkú pozornosť na priemyselnej aj akademickej pôjde. Výskumníci a prevádzkovatelia sietí pracujú ako aj s drôtovými pevnými sieťami, tak aj bezdrôtovými sieťami a problémami s nimi spojenými. Kvôli rozmanitosti a zložitosti sietí sú sieťové algoritmy navrhnuté na špecifické scenáre, charakteristiky siete a požiadavky používateľov. Tieto problémy by mohli byť vyriešené už spomínaným strojovým učením, ktoré už v iných oblastiach ako sú bio-informatika, rozpoznávanie reči dosiahli značné výsledky [2].

Strojové učenie sa snaží konštruovať algoritmy a modely, ktoré sa môžu naučiť rozhodovať priamo z údajov bez dodržiavania vopred definovaných pravidiel. Riešenie zložitých problémov je jedným z najväčších výhod strojového učenia. Pri niektorých úlohách vyžadujúcich klasifikáciu, regresiu a rozhodovanie môže strojové učenie dosiahnuť lepšie výsledky ako ľudské bytosti. Vzhľadom na to, že sieťové polia často riešia komplexné problémy, ktoré potrebujú čo najefektívnejšie riešenie, je využitie strojového učenia veľmi sľubné riešenie. Začlenenie strojového učenia pri návrhu a spravovaní sietí tiež poskytuje možnosť generovania nových sieťových aplikácií. Pokrok v oblasti možného počítačového výkonu ako sú GPU a TPU, knižnice strojového učenia ako Tensorflow a Scikit-Learn poskytujú príležitosť na plné využitie potenciálu strojového učenia v sieťových systémoch.



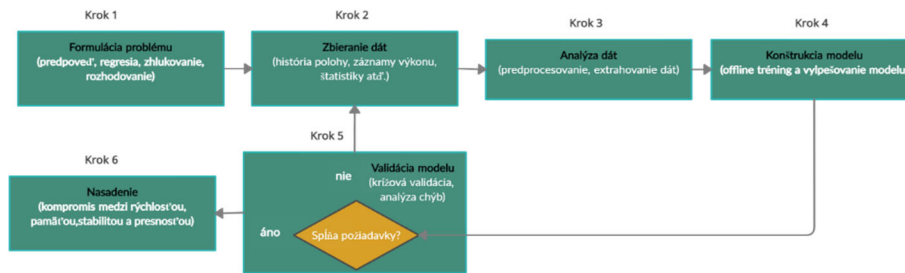
Obr. 2 Základná idea strojového učenia pre MANET siete.

Strojové učenie anglicky Machine Learning (ML) je odvetvie umelej inteligencie, ktoré umožňuje počítačovým systémom učiť sa priamo z príkladov, údajov a skúseností. Prostredníctvom umožnenia počítačom inteligentne vykonávať konkrétne úlohy, strojovo vzdelávacie systémy môžu vykonávať komplexné procesy pomocou samo učenia sa z dát bez nutnosti vopred naprogramovaných pravidiel. Cieľom strojového učenia je zostaviť počítačové programy, ktoré sa môžu učiť z veľkého množstva údajov. Tento postup možno porovnať s prístupom umelej inteligencie, avšak strojové učenia sa zaoberá oveľa širším poľom vedeckých disciplín ako sú štatistika, kognitívna veda a informačné technológie. Strojové učenie sa stalo jednou z najdôležitejších tém v rámci rozvojových organizácií, ktoré hľadajú inovatívne riešenia a spôsoby, ako využiť dáta a pomôcť tak podniku získať novú úroveň porozumenia. Keďže nové dáta sú neustále pridávané, modely strojového učenia zabezpečujú, že riešenia sú taktiež neustále aktualizované. Ak sa pri strojovom učení použijú vhodné dáta, ktoré sú ešte k tomu neustále aktualizované, model je schopný do určitej miery predvídať budúce eventy. Avšak strojové učenie nie je jednoduchý proces. Strojové učenie využíva rôzne algoritmy, ktoré sa opakovaním učia z údajov za účelom neustáleho zlepšovania sa v popisovaní údajov a predvídaní výsledkov. Keďže tieto algoritmy, taktiež prijímajú údaje o ich tréningu, je možné na základe týchto údajov vytvoriť ešte presnejšie modely. Model strojového učenia je výstup generovaný pri tréningu stroja pomocou daných dát. Po ukončení tejto tréningovej fázy budete schopní vložiť do modelu určité vstupné dáta a na základe toho dostať určité výstupné dáta. Prístupy ML možno rozlíšiť z hľadiska reprezentácie a adaptácie. Systém ML musí ukladať naučené informácie do dátovej štruktúry, ktorá je nazývaná hypotéza a má formu modelu. Cieľom tejto hypotézy je zovšeobecniť zadané údaje a uprednostniť najjednoduchšiu hypotézu. Pravidlom je, že hypotéza by mala byť jednoduchšia ako samotné údaje. Algoritmy učenia určujú, ako aktualizovať naučenú hypotézu a skúsenosti, aby sa výkonnosť merala s ohľadom na optimalizáciu úloh [3].

Využitie strojového učenia pri MANET (MLN) je o to viac výhodnejšie ako v iných oblastiach z nasledujúcich dôvodov. Po prvé, jedny z najväčších výhod strojového učenia sú klasifikácia a predikcia. Práve tieto schopnosti hrajú základnú, ale dôležitú rolu pri úlohách a riešeníach problémov v MANET sieťach, ktoré sú napríklad detekcia narušenia, alebo predpoveď výkonu. Strojové učenie môže uľahčiť pri rozhodovanom procese uzlov, plánovaní infraštruktúry siete a prispôbení daných parametrov. Po druhé, rôzne sieťové problémy potrebujú komunikovať s veľmi komplikovanými systémovými prostrediami. Nie je jednoduché vybudovať presné alebo analytické modely, ktoré reprezentujú komplexné systémové správanie ako je napríklad rozloženie zaťaženia, alebo priepustnosť siete. Strojové učenie avšak môže poskytnúť odhadovaný model pre takéto systémy s prijateľnou presnosťou. Nakoniec, každý sieťový scenár môže mať rôzne charakteristiky a výskumníci musia riešiť problémy pre každý tento scenár samostatne. Strojové učenie im môže poskytnúť nové možnosti ako tieto problémy riešiť, môže zostrojiť všeobecný model prostredníctvom tréningových metód, ktorý sa potom pozmení a aplikuje na konkrétne problémy [4], [5].

Na obr. 3 môžeme vidieť základnú blokovú schému toho ako funguje strojové učenie v MANET. Jednotlivé etapy sú nezávislé, ale majú definované pevné vnútorné vzťahy. Táto bloková schéma je podobná s tradičnou blokovou schémou strojového učenia, pretože problémy v MANET sieťach sú dátovo podobné tým v iných aplikáciách. Samotné strojové učenie

pozostáva z nasledovných etáp [4], [5], [6]:

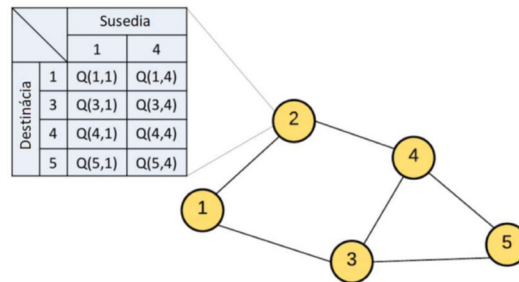


Obr. 3 Základný model strojového učenia pre MANET siete.

- **Formulácia problému:** Pretože tréningový proces strojového učenia je často časovo náročný a vyžaduje vysoké náklady, je dôležité správne formulovať problém. Cieľový problém je možné klasifikovať do jednej z kategórií strojového učenia, napr. klasifikácia, zoskupovanie alebo rozhodovanie. Toto kategorizovanie pomáha trénovanému modelu rozhodovať koľko a aké dáta sa k nemu dostávajú a ktoré z nich sú pre neho najdôležitejšie. Nesprávne formulovanie cieľového problému môže viesť k vytvoreniu nevhodného modelu učenia, ktorého výsledkom môže byť neuspokojivá presnosť v učení.
- **Zber údajov:** Cieľom tohto kroku je zhromaždiť veľké množstvo reprezentatívnych sieťových údajov bez zaujatosti. Sieťové údaje napr. stopy z komunikácie, komunikačné logy s údajmi o prenose a výkone sa zaznamenávajú z rôznych sieťových vrstiev podľa potreby aplikácie. Napríklad problém s klasifikáciou prenosu často vyžaduje dátové súbory obsahujúce stopy na úrovni paketov označené príslušnými triedami. V kontexte MLN sa údaje často zhromažďujú v dvoch fázach. V offline fáze je najdôležitejšie zhromaždenie dostatku kvalitných historických údajov, ktoré sú potrebné pre analýzu a výcvik modelov. V online fáze sa často ako vstup využívajú štatistiky v reálnom čase siete pre učiaci sa model. Novo zhromaždené dáta možno môžu byť použité na nahradenie starých dát, kvôli adaptácii modelu a udržaniu presnosti.
- **Analýza údajov:** Každý sieťový problém má svoje vlastné charakteristiky a je ovplyvnený mnohými faktormi. Z týchto faktorov, ale len niekoľko má výrazný vplyv na výkon siete. Napríklad oneskorenie paketov (RTT) a čas príchodu (ACK) môžu byť kritickými vlnosťami pri výbere najlepšej veľkosti TCP okna. V paradigme učenia je nájdenie správnych funkcií kľúčom k čo najefektívnejšiemu využitiu údajov. Tento krok sa pokúša získať efektívne vlastnosti sieťového problému analýzou vzoriek historických údajov. Pred extrakciou údajov funkcií je dôležité predbežne spracovať a vyčistiť nespracované údaje prostredníctvom procesov, ako je normalizácia, diskreditácia a doplnenie chýbajúcich hodnôt.
- **Konštrukcia modelu:** Zahŕňa výber modelu, jeho školenie a ladenie. Je potrebné zvoliť vhodný model učenia podľa veľkosti údajov, ich typických charakteristík a kategórií problému. Napríklad presná predikcia priepustnosti môže zlepšiť adaptáciu bitového toku internetového videa.
- **Validácia modelu:** Offline validácia je nevyhnutný krok v pracovnom postupe MLN na vyhodnotenie toho, či výučbový algoritmus funguje s uspokojivou presnosťou. Počas tohto kroku sa na overenie celkovej presnosti modelu zvyčajne používa krížová validácia, aby sa zistilo, či je model over-fitng alebo under-fitting, čiže či model nedosiahol, alebo presiahol požadované charakteristiky. To poskytuje dobrý náhľad na to, ako optimalizovať model, napríklad zvýšením objemu dát, alebo znížením zložitosti modelu. Analýza nesprávnych vzoriek pomáha nájsť príčiny chýb.
- **Nasadenie a odvodenie:** Pri implementácii modelu výučby v pracovnom prostredí musia byť zvažované určité problémy týkajúce sa výpočtovej sily a energetických zdrojov, požiadavkám na čas odozvy, prenosovú rýchlosť atď. Pre optimálny a čo najefektívnejší výsledok je preto nutné vytvoriť kompromis medzi prenosom a zaplňaním siete paketami. Okrem toho strojové učenie pracuje stále k dosiahnutiu čo najlepšieho úsilia a neposkytuje žiadnu záruku výkonu, čo vyžaduje, aby návrhári systému brali do úvahy odolnosť voči možným chybám. Nakoniec praktické aplikácie často vyžadujú, aby model pracoval s vstupnými údajmi v reálnom čase.

### A. Smerovací algoritmus na báze Q- učenia v MANET

Medzi prvé smerovacie protokoly založené na RL učení bol protokol Q-Smerovanie (Q-Routing), ktorý využíva Q-učenie. Hodnoty Q sa ukladajú do tabuľky so záznamom pre každý práve využívaný prenos medzi uzlami aj s jeho dosiahnuteľnými susedmi. Na základe funkcie odmeňovania sa algoritmy RL pokúšajú nájsť trasy s čo najmenším oneskorením. Ide o techniku učenia formou odmeňovania (RL) v ktorej má softvérový agent iba obmedzené znalosti prostredia v ktorom sa nachádza, čo vedie k vysokej miere neistoty toho ako bude prostredie reagovať na vykonané akcie. Interakcie s prostredím sú jediným spôsobom, akým sa agent môže učiť. V každom kroku agent dostane určitú veľkosť odmeny, ktorá slúži ako feedback na jeho vykonané akcie. Agentov gól je naučiť sa, ktoré akcie spraviť v ktorom kroku, aby maximalizoval odmenu. Pomocou Q-učenia sa poznatky týkajúce veľkosti odmien a stavu prostredia získavajú prostredníctvom interakcie s týmto prostredím (Obr. 4) [6].



Obr. 4 Princíp činnosti Q-učenia v MANET sieti.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (1)$$

V Q-učení sa Q funkcia používa na meranie kvality vykonanej akcie pomocou prijatých odmien a jej učenia (1). Smerovací algoritmus si v definovanom čase  $t$  aktualizuje hodnotu  $Q(s_t, a_t)$  po tom, ako agent vykoná  $a_t$  akciu, prostredníctvom ktorej sa zmení stav  $s_t$  na  $s_{t+1}$  a následne dosiahne odmenu  $r_{t+1}$ .  $\max_a Q(s_{t+1}, a)$  je očakávaná maximálna hodnota Q-funkcie, ktorý dosiahne agent pri výbere  $a$  činnosti v nasledujúcom stave  $s_{t+1}$ .

### B. Modifikácie algoritmu Q- učenia v MANET

Rozšírenie s názvom Prediktívne Q-Smerovanie (PQ-Smerovanie) bolo navrhnuté na riešenie problému chýbajúceho prieskumu pôvodného Q-smerovania použitím zlomku z paketov na objavovanie nových trás. Tento prieskum sa vykonáva na základe minulých skúseností zaznamenaných na skúmaných trasách. Uzly ukladajú najlepšie Q hodnoty pre každú trasu, ako aj najnovšie Q hodnoty a odhadovanú hodnotu doby obnovovania spojenia. Tieto hodnoty sa potom použijú na riadenia dopravy objavovacích paketov. Podobne ako Q-Smerovanie aj PQ-Smerovanie používa smerovaciu politiku, ktorá sa rozhoduje na základe Q hodnôt, ale sú ešte ďalej vylepšené o informácie dĺžok obnovenia spojenia. Výsledky ukázali, že PQ-Smerovanie funguje lepšie ako Q-Smerovanie [5], [7].

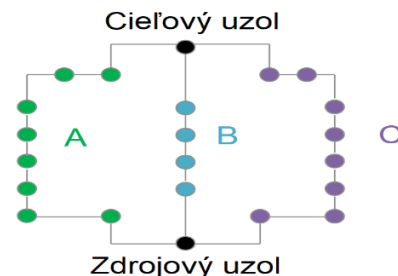
Ďalším smerovacím protokolom pre MANET siete založeným na princípe Q-Učenia je Reinforcement Learning Based Geographic Routing Protocol (RLGR). RLGR bol navrhnutý pre senzorové siete, kde uzly majú prístup k informáciám o svojej polohe a môžu tieto informácie vymeniť s ďalšími uzlami. V RLGR je životnosť batérie najdôležitejšou metrikou, ktorú je nutné optimalizovať. V prípade, že uzlu dochádza batéria, alebo nemá prístup k hlavnému toku paketov v sieti, uzol odošle naspäť negatívnu odmenu na okolité uzly. Naopak ak je uzol blízko, alebo má prístup k tomuto hlavnému toku výmeny informácií odošle okolitým uzlom pozitívne odmeny. Na komunikáciu sa inak využíva klasický výpočet metriky pozostávajúci s oneskorenia a vzdialenosti. Na prieskum nových trás a zabraňovaniu vytváraniu slučiek RLGR využíva špeciálne upravené pakety [7], [8].

Môžeme si všimnúť, že zatiaľ čo tradičné MANET smerovacie algoritmy nevyužívajú všetky potrebné vlastnosti dynamických ad hoc sietí v svoj prospech, tak na druhú stranu hybridné smerovacie algoritmy založené na strojovom učení nevyužívajú na plno objavovanie nových trás a tým nemôžu garantovať nájdenie optimálnych trás. Súvisiacim problémom sú algoritmy, ktoré síce využívajú oba tieto prvky, ale za cenu výrazného zväčšenia dátového prenosu ako napríklad v RLGR. Tento dátový prenos je navyše smerovaný neznámymi cestami s neoptimálnymi metrikami, čo zvyšuje pravdepodobnosť už tak dosť existujúcej hrozby straty paketov a potrebu opätovného prenosu. Výmena informácií v súčasných protokoloch založených na učení formou

odmeňovania často využíva broadcast, čo je vysoko neefektívne a spôsobuje zahlcovanie siete a ešte väčšiu spotrebu energie [8].

## II. DOSIAHNUTÉ VÝSLEDKY SIMULÁCIÍ

Za účelom analyzovania a implementovania jednotlivých metód algoritmov strojového učenia pre MANET siete sme úspešne implementovali AODV, Q - smerovací protokol a Q<sup>2</sup> smerovací protokol do simulačného nástroja NS-3. Na porovnanie nami implementovaných algoritmov medzi sebou, sme vytvorili testovaciu sieť. Nakoľko chceme vidieť, ktorý z protokolov podal aký výkon za rovnakých podmienok, musia byť tieto protokoly testované na rovnakej sieti s rovnakými scenármi eventov.



Obr. 5 NS-3 – grafické znázornenie rozloženie MANET uzlov.

Experimenty sa teda uskutočňovali na MANET sieti tvorenej z 22 mobilných uzlov. Tieto uzly sme rozložili na tri hlavné cesty s predefinovaným zdrojom a cieľom komunikácie. Grafické zobrazenie siete môžeme vidieť na obr. 5. Dve krajné cesty A,B majú dĺžku 9 skokov a stredná cesta B má dĺžku 5 skokov. Experimenty boli prevedené pomocou zmeny úrovne straty paketov, oneskorenia paketov, zmenou úrovne kolísania oneskorenia (jitter). Kolísanie oneskorenia je pomenované pre oneskorenie príchodu paketov, spôsobených preťažením siete alebo zmenami v topológii. Štandardná dĺžka kolísania oneskorenia je pár milisekúnd (ms), ale oneskorenie väčšie než 10ms môže viesť k strate paketov (Tab. 1).

Protokoly, ktoré sme si vybrali boli AODV, Q smerovací protokol a Q<sup>2</sup> smerovací protokol. Dôvod prečo sme si vybrali AODV, bol z dôvodu, že AODV tvorí jadro týchto protokolov a môžeme pozorovať ako implementácia strojového učenia zlepšila výsledky v smerovaní. Zároveň ide o široko používaný dynamický protokol, ktoré je využívaný aj v reálnom svete. Pre všetky experimenty sme použili diskontný faktor  $\gamma$  rovný 1 a faktor rýchlosti učenia  $\alpha = 0.5$ . Tieto parametre sa ukázali, ako dobrý kompromis medzi rýchlou konvergenciou v sieti (reagovanie na zmeny v topológii) a úrovňou zahltenia siete.

Výsledky v tabuľke č. 2 a obr. č. 6 ukazujú na to, ako Q<sup>2</sup> - smerovací protokol zmení rýchlejšiu trasu B na pomalšiu trasu C, kvôli 66% strate paketov v čase  $t = 1000$ . AODV spravilo to isté z dôvodu počtu nedoručených Hello paketov. Zo všetkých simulácií išlo o jediný prípad, kedy AODV bolo schopné podať rovnaký výkon ako Q<sup>2</sup> smerovací protokol využívajúci QoS. Pri ďalšom pridaní oneskorenia na trasu B, zareagoval aj Q smerovací algoritmus, ktorý zmenil trasu na C. V čase  $t = 3000$  sme pridali na trasu C ďalšiu 27% stratu paketov, na čo Q<sup>2</sup> smerovací protokol hneď zareagoval a prešiel z používania trasy C na pomalšiu trasu A. AODV eventuálne spravilo to isté, ale ešte predtým na istý čas prešiel na trasu B, na ktorom bola stále nastavená strata paketov. Q smerovací protokol nezareagoval vôbec. Na konci simulácie v čase  $t = 4500$  sme odstránili spomalenie aj stratu paketov z trasy B, čím sme z nej vytvorili novú najoptimálnejšiu trasu.

Tabuľka 1  
 Nastavenie hodnôt pre proces smerovania

Čas [s]	Cesta	Event	Hodnota
0	C	Oneskorenie paketov	19 ms
0	B	Oneskorenie paketov	1 ms
0	A	Oneskorenie paketov	31 ms
1000	B	Strata paketov	66 %
2000	B	Oneskorenie paketov	60 ms
3000	C	Strata paketov	27 %
4500	B	Oneskorenie paketov	3 ms
4500	B	Strata paketov	0 %

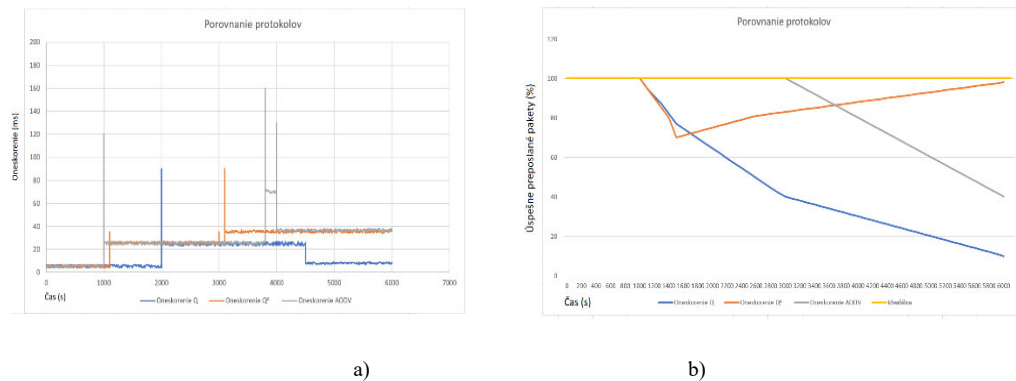
Q<sup>2</sup> smerovací protokol avšak nezmenil komunikáciu na trasu B. Toto bolo zapríčinené neobnovením metriky trasy B, kde Q<sup>2</sup> smerovací protokol si stále myslel, že na trase prebieha strata paketov. Q smerovací protokol avšak zareagoval a zmenil optimálnu trasu na B.

Pomer úspešne preposlaných paketov odoslaných za daný čas simulácie ukazuje nasledujúce. Na obr. môžeme vidieť, ako Q<sup>2</sup>-Smerovací protokol stratil úspešne preposlané pakety okolo času

$t = 1000$  kvôli nášmu eventu straty paketov, ale hneď na to sa zotavil zo straty a pri druhom evente straty paketov v čase  $t = 3000$  už nebol zasiahnutý. Na druhú stranu Q smerovací protokol zaznamenal obrovskú stratu úspešne preposlaných paketov a nebol schopný dobehnúť túto stratu. A protokol AODV bol zasiahnutý až druhou vlnou straty paketov.

Tabuľka 2  
 Výber komunikačných ciest pre jednotlivé smerovacie protokoly v MANET

Čas [s]	AODV	Q smerovanie	Q <sup>2</sup> smerovanie
0	B	B	B
1000	C	B	C
2000	C	C	C
3000	B	C	A
4500	I	2	I



Obr. 6 Oneskorenie doručovania paketov (a) a úspešné preposielanie paketov (b) - analýza dosiahnutých výsledkov ML smerovacích algoritmov pre MANET.

### III. ZÁVER

V práci sme analyzovali možnosti využitia strojového učenia v procese smerovania v MANET sieťach. Následne sme analyzovali jednotlivé metódy a algoritmy strojového učenia vhodné pre MANET siete. Išlo o porovnanie troch sieťových protokolov navrhnutých pre MANET siete, z týchto troch protokolov dva vyžívajú pri smerovaní rôzne technicky strojového učenia. Ide o protokoly Q smerovanie, Q<sup>2</sup> smerovanie. získané výsledky a potvrdili sme, že využitie strojového učenia v smerovacích protokoloch v procese smerovania vedie k zlepšeniu smerovacích aspektov týchto sietí, ako je menšia strata a oneskorenie paketov, lepší reakčný čas na zmeny topológie v sieti a výber vhodnejšej trasy. Z nami testovaných protokolov Q<sup>2</sup> predstavuje najlepšiu metódu využitia strojového učenia pri procese smerovania v MANET sieťach.

### POĎAKOVANIE

Táto publikácia vznikla vďaka podpore Agentúry na podporu výskumu a vývoja v rámci projektu c. APVV-17-0208 - Odolné mobilné siete na doručovanie obsahu.

### REFERENCIE

- [1] P. Bellavista, G. Cardone, A. Corradi and L. Foschini, "Convergence of MANET and WSN in IoT urban scenarios", IEEE Sensors, vol. 13, no. 10, pp. 3558-3567, 2013.
- [2] R. Bruzgiene, L. Narbutaite, T. Adomkus, "MANET Network in Internet of Things System", InTech, 2017, 10.5772/66408
- [3] R. J. Cai, X. J. Li and P. H. J. Chong, "An evolutionary self-cooperative trust scheme against routing disruptions in MANETs", IEEE Trans. Mobile Comput., vol. 18, no. 1, pp. 42-55, Jan. 2019.
- [4] S. Kanthimathi, J. R. Prathuri, "Classification of Misbehaving nodes in MANETS using Machine Learning Techniques", 2nd PhD Colloquium on Ethically Driven Innovation and Technology for Society (PhD EDITS), 2020, pp. 1-2, doi: 10.1109/PhDEDITS51180.2020.9315311.
- [5] R. Akbani, T. Korkmaz, G.V.S. Raju, "ML Based Reputation System for Defending Against Malicious Node Behavior", IEEE Explore, 2018.
- [6] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, "Machine learning for wireless communication in the Internet of Things: A comprehensive", in Ad Hoc Networks, vol. 93, 2019.
- [7] K. S. Sankaran, N. Vasudevan, K. R. Devabalaji, T. S. Babu, H. H. Alhelou and T. Yuvaraj, "A Recurrent Reward Based Learning Technique for Secure Neighbor Selection in Mobile AD-HOC Networks," in IEEE Access, vol. 9, pp. 21735-21745, 2021, doi: 10.1109/ACCESS.2021.3055422.
- [8] D. Ma, X. Hu, H. Zhang, Q. Sun and X. Xie, "A hierarchical event detection method based on spectral theory of multidimensional matrix for power system", IEEE Trans. Syst. Man Cybern. Syst., Aug. 2019.

# Bezpečnosť a Blockchain v MANET

<sup>1</sup>Michaela ŠAKOVÁ, <sup>2</sup>Ján PAPAJ

<sup>1 2</sup> Katedra elektroniky a multimediálnych technológií,  
Fakulta elektrotechniky a informatiky Technická Univerzita Košice,  
Slovenská republika

<sup>1</sup>michaela.sakova@student.tuke.sk, <sup>2</sup>jan.papaj@tuke.sk

**Abstract** — Článok sa venuje problematike využitia technológie Blockchain vo viacpreskokových mobilných sieťach (MANET) pohľadu sieťovej bezpečnosti. Rýchly rozvoj dátovej komunikácie si vyžaduje aplikovanie alternatívnych prístupov na vytvorenie spoľahlivej a bezpečnej komunikačnej cesty medzi mobilnými terminálmi v MANET sieti. Práve technológia Blockchain sa javí ako efektívny a progresívny nástroj na zvýšenie sieťovej bezpečnosti v MANET s cieľom na eliminovanie zlomyseľných a nebezpečných uzlov v MANET.

**Keywords** — MANET, smerovanie, bezpečnosť, Blockchain

## I. ÚVOD

Bezpečnosť v mobilnej viacpreskokej komunikácii, ktorá je formovaná dynamicky sa pohybujúcimi uzlami, je dôležitou súčasťou jej základnej funkčnosti. Dostupnosť sieťových služieb, dôveryhodnosť a integrita údajov sa dá dosiahnuť splnením mnohých bezpečnostných otázok. Mobilné viacpreskokové siete (MANET) sú často postihované niekoľkými bezpečnostnými útokmi v dôsledku svojich vlastností, ktorými sú [1], [2], [3]:

- *otvorené médium,*
- *dynamicky meniacia sa topológia,*
- *kooperatívne algoritmy,*
- *nedostatok centrálného monitorovania a riadenia,*
- *žiadny jasný obranný mechanizmus.*

Tieto faktory zmenili situáciu v boji proti bezpečnostným hrozbám. V posledných niekoľkých rokoch sa veľa diskutovalo o zabezpečení počítačových sietí. Väčšina diskusií sa týkala iba statiky a sietí založených na drôtových systémoch. Mobilné multihop siete si však naďalej vyžadujú ďalšie diskusie a vývoj v oblasti bezpečnosti.

S neustálym nástupom nových prístupov k vytváraniu sietí sa objavujú nové problémy týkajúce sa základov smerovania. Smerovacie protokoly určené primárne pre internet sa líšia od mobilných multihop sietí. Tradičná smerovacia tabuľka bola v podstate vyrobená pre hostov, ktorí sú pripojení káblovým spojením k nedynamickému chrbtvej kosti. Vzhľadom k tomu nie je možné podporovať mobilné multihop siete kvôli pohybu komunikačných uzlov a dynamickej topológii sietí. Z dôvodu rôznych faktorov vrátane nedostatku pevnej infraštruktúry, absencie už vytvoreného dôverného vzťahu medzi rôznymi uzlami a dynamickej topológii sú smerovacie protokoly zraniteľné voči rôznym útokom.

Hlavnými zraniteľnosťami, ktoré boli doteraz preskúvané, sú typy útokov, ktoré zahŕňajú sebectvo, dynamický charakter, prísne obmedzenie zdrojov a tiež otvorené sieťové médium. Mobilné multihop siete pracujú bez centralizovanej správy, kde uzly medzi sebou komunikujú na základe vzájomnej dôvery. Táto vlastnosť ich robí náchylnejšími na napadnutie útočníkom z vnútra siete.

Aj z dôvodu bezdrôtových spojení sú mobilné multihop siete náchylnejšie na útoky. Všetky spomínané okolnosti spôsobujú to, že útočník ľahko vstupuje do siete a získava prístup k prebiehajúcej komunikácii. Mobilné uzly, ktoré sa nachádzajú v dosahu bezdrôtového spojenia, môžu sieť odpočúvať a dokonca sa jej môžu zúčastniť [4].

## II. TYPY ÚTOKOV V MANET

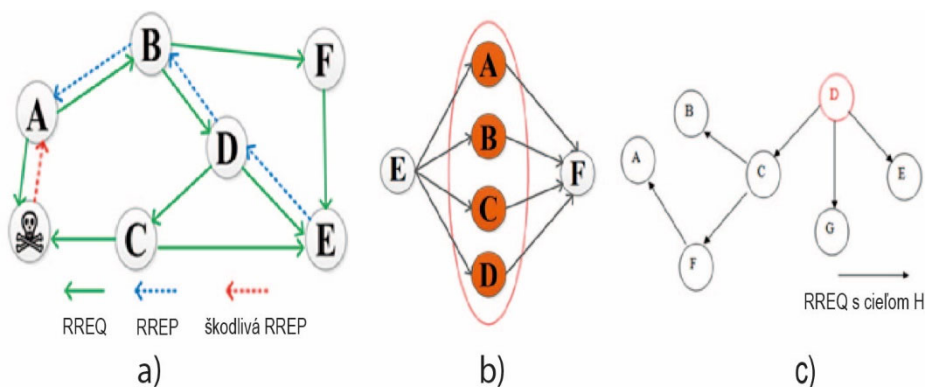
Útoky v MANET možno všeobecne klasifikovať ako pasívne a aktívne útoky [4], [5], [6].



- *Pasívne útoky* - Pri týchto útokoch útočník vôbec nezasahuje do sieťovej premávky. Odpočúvaním sa snaží získať čo najviac informácii, ktoré sa prenášajú medzi dvoma stranami. Pasívny útočník počúva kanál, a tak pakety obsahujúce tajné informácie (napr. IP adresy, umiestnenie uzlov atď.) môžu byť odcudzené, čo porušuje paradigmu dôvernosti [6].
- *Aktívne útoky* - zatiaľ čo pasívne útoky nemajú za cieľ narušiť sieťové operácie, aktívne útoky sa pokúšajú upraviť alebo zmeniť údaje vymieňané v sieti, čo môže narušiť normálne fungovanie sietí. Pri aktívnych útokoch môžu útočníci modifikovať, vložiť, odhodiť pakety alebo na spustenie útoku môžu použiť rôzne funkcie siete [5],[6].

Útoky je tiež možné klasifikovať v závislosti od polohy útočníka v sieti.

- *Externé útoky* - sú útoky neoprávnených uzlov, ktoré nie sú súčasťou siete. Cieľom týchto útokov je zvyčajne spôsobiť preťaženie siete, odoprieť prístup k určitým sieťovým funkciám alebo narušiť celú sieťovú prevádzku [6].
- *Interné útoky* - vykonávajú autorizované uzly v sieti. Dôvodom ich škodlivého správania môže byť:
  - únos týchto (autorizovaných) uzlov nejakým externým útočníkom a ich následné použitie na uskutočnenie interných sieťových útokov,
  - sebeckvo šetriť svoje obmedzené zdroje, ako je kapacita batérie, možnosti spracovania a šírka pásma komunikácie, a využívanie ďalších uzlov v ich prospech.



Obr. 1 Příklad útokov v procese smerovania v MANET: a) Black hole, b) Sybil, c) Flooding.

Keďže útočníci sú už súčasťou siete ako autorizované uzly, vnútorné útoky sú v porovnaní s externými útokmi závažnejšie a je ťažšie ich odhaliť [7]. Jednou z vážnych hrozieb MANET je smerovací útok, pri ktorom môže narušený uzol pozdĺž smerovacej cesty zhodiť alebo upraviť pakety pri ich preposielaní.

#### A. Útok typu čierna diera

Čierna diera (Black hole) označuje oblasť v sieti, ktorá zhadzuje dopravu smerujúcu do konkrétneho cieľa, bez informovania zdrojového uzla o úspešnom doručení paketov do cieľa. Pri tomto útoku uzol využíva smerovací protokol, aby sa prejavil ako uzol, ktorý má najkratšiu cestu ku konkrétnemu cieľu. Tento uzol získa dátové pakety, ktoré ním majú byť odoslané ďalej do určeného cieľa. Uzol tieto pakety zahodí, spotrebuje ich alebo využije svoje umiestnenie v sieti na vlastnú inzerciu ako cieľový uzol a začne presmerovať rôzne pakety vo vnútri siete. V takom prípade zdroj a cieľ nemôžu navzájom komunikovať (Obr. 1a) [7].

Uzol A chce poselať údaje do uzla E. Vysiela preto správu RREQ, ktorá sa šíri sieťou tak, ako je to uvedené na [8]. Škodlivý uzol odpovedá falošnou správou RREP, čo značí, že má najkratšiu cestu do cieľa. Táto falošná správa dorazí do uzla A, ktorý následne predpokladá, že proces zisťovania trasy sa skončil. Uzol A začne poselať pakety, určené uzlu E, škodlivému uzlu. Škodlivý uzol tieto pakety zahodí a spôsobí nedostupnosť uzla E pre uzol A, alebo sám inicializuje proces hľadania cesty do uzla E, aby mohol deformovať komunikáciu medzi týmito uzlami. Útok čiernej diery je možné klasifikovať na základe stratégie použitej škodlivým uzlom na vykonanie útoku. Uzol buď odhodí všetky pakety, ktoré k nemu dorazia a majú byť presmerované na určené miesto, alebo niektoré pakety zahodí.

Uzol, ktorý plánuje zaútočiť, musí nájsť spôsob, ako sa dostať do premávky, v ktorej sa budú

doručovať kontrolné alebo dátové pakety. Spoliehajú sa na niektoré zraniteľnosti v použitom smerovacom protokole, ktorý bol navrhnutý na základe dôveryhodnosti medzi sieťovými uzlami, môže každý uzol urobiť nesprávne správanie a sabotovať sieťové operácie zničením dátových paketov alebo zneužitím riadiacich paketov. Zahodením paketov sa ukončí komunikácia a prenos medzi dvoma uzlami, čo je však horšie - škodlivý uzol bráni vytvoreniu trasy medzi týmito uzlami.

### B. Sybil útok

Škodlivé uzly v sieti sa dokážu vydávať nielen za jeden uzol, ale môžu tiež predstierať identitu viacerých uzlov, čím sa oslabuje redundancia mnohých smerovacích protokolov. Tento útok sa nazýva Sybil (Obr. 1b). Útok Sybil sa prejavuje fíngovaním viacerých identít škodlivého uzla predstieraním, že sa skladá z viacerých uzlov v sieti. Jediný uzol teda môže prevziať úlohu viacerých uzlov a môže monitorovať alebo brániť viacerým uzlom súčasne. Ak sa ten istý paket pošle cez niekoľko odlišných ciest, možno ľahko zistiť zmenu v paketoch prichádzajúcich z jednej z týchto ciest. Takto je možné izolovať votrelca v sieti. Ak je však jeden nebezpečný uzol schopný zastupovať niekoľko ďalších uzlov, účinnosť týchto opatrení sa výrazne zníži. Útočník môže získať prístup ku všetkým fragmentovaným informáciám alebo môže zmeniť všetky pakety v rovnakom prenose, takže cieľový uzol už nebude môcť zistiť neoprávnenú manipuláciu [7], [8].

Útok Sybil je zameraný najmä na prostredia distribuovaných systémov. Sybilský uzol si môže vyrobiť novú identitu sám pre seba alebo ukradne identitu legitímneho uzla. Niektoré efekty spôsobené útokmi Sybil sú:

- *v prítomnosti sybilských uzlov v sieti môže byť ťažké identifikovať zlomyselný uzol,*
- *útoky Sybil, bránia spravodlivému pridelovaniu zdrojov medzi uzlami v sieti,*
- *v určitých aplikáciách môžu byť senzory použité na vykonávanie hlasovania pri rozhodovaní. Výsledok hlasovacieho procesu sa môže líšiť v dôsledku prítomnosti duplicitných identít,*
- *uzly Sybil ovplyvňujú normálnu činnosť smerovacích protokolov tým, že sa objavujú na rôznych miestach v sieti.*

Na Obr. 1b uzol D generuje pakety RREQ určené na adresu uzla H, ktorý v skutočnosti v sieti neexistuje, a vysielá ich do všetkých susedných uzlov (C, G, E). Pretože žiadny susedný uzol nebude schopný nájsť uzol H, znova ho vysielajú za predpokladu, že niektoré ďalšie uzly budú schopné nájsť cestu k uzlu H [8].

### C. Záplavový útok

Tento útok je ľahké realizovať, no dokáže spôsobiť najviac škôd. Dosahuje sa zahltením správami RREQ alebo falošnými dátami. Najčastejšou záplavou na sieťovej vrstve je záplava RREQ. Je to DoS útok, pri ktorom škodlivé uzly využívajú proces zisťovania trasy reaktívnych smerovacích protokolov.

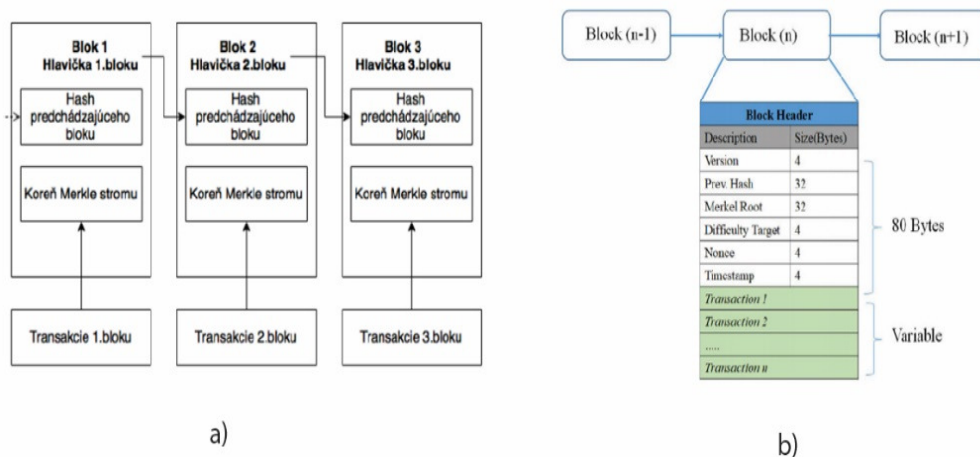
Pri tomto útoku škodlivý uzol generuje veľké množstvo správ RREQ a vysielá ich do neplatných cieľov. Keď sú takéto falošné pakety RREQ vysielané do siete vo vysokom počte, sieť je nimi presýtená a nie je schopná prenášať dátové pakety. Vedie to teda k preťaženiu siete. Na uzol E môže komunikovať s uzlom F cez 4 rôzne uzly. Avšak tieto uzly (A, B, C, D) sú len identitami jedného škodlivého uzla [8], [9].

## III. BLOCKCHAIN V MANET

Blockchain ako distribuovaná databáza uchováva neustále rozširujúci sa počet časovo usporiadaných záznamov, známych tiež ako transakcie. Transakcie sú zoskupené do blokov a tvoria kryptografický hašový reťazec (HASH). Táto databáza sa stala široko popularizovanou potom, čo bola použitá ako základná technológia bitcoinovej kryptomeny [10]. Je implementovaná ako spojovaný zoznam (linked list). Položky zoznamu sú označované ako bloky. Previazanosť medzi jednotlivými blokmi je realizovaná pomocou hašov aktuálneho a predchádzajúceho bloku. Bloky v blockchaine sú usporiadané podľa času ich vzniku. Každý z blokov potvrdzuje transakcie vykonané v predchádzajúcich blokoch. Pomocou Obr. 5 je ilustrovaná schéma blockchainu [11]. Prvý blok v blockchaine je označovaný ako Genesis blok [19].

Každý blok obsahuje hlavičku a telo. Hlavička bloku pozostáva z hašovacej funkcie predchádzajúceho bloku, hodnoty nonce, časovej značky predchádzajúceho bloku a Merkle root hašu (hašu všetkých transakcií v bloku), ako je znázornené na Obr. 5a. Telo bloku obsahuje

zoznamy transakcií a niektoré ďalšie údaje, v závislosti od požiadavky blockchainu. Reťazec blokov je zdieľaný naprieč všetký zúčastnenými uzly (Obr. 5b).



Obr. 2 Blockchain a MANET: a) ukážka štruktúry Blockchainu pre 3 uzly, b) hlavička Blockchainu.

Nové bloky sú overované a potvrdzované ostatnými zúčastnenými uzlami, čo eliminuje potrebu centrálného vedenia. Pripojenie nového bloku k reťazcu (nazývané tiež ako ťažba) znamená riešenie výpočtovo náročnej, ťažko riešiteľnej a ľahko overiteľnej kryptomeny. Jednotlivé uzly sa musia dohodnúť na stave blockchainu, čo znamená, že musia dosiahnuť konsenzus o tom, ktoré transakcie sa uskutočnili v akom poradí. Rozlišujeme dva hlavné typy blockchainov [13]:

- *verejný blockchain (bez povolenia),*
- *súkromný blockchain (s povolením).*

Verejné blockchajny sú otvorené, voľne prístupné. Tým pádom sa k nim môže ktokoľvek pripojiť, prispieť do nich, alebo zobrazit' ich obsah bez potreby povolenia od centrálnej autority. Predstavujú skutočnú decentralizáciu a transparentnosť, avšak ich obsluha a údržba je spravidla pomalšia a nákladnejšia, pretože je potrebné použiť zložitejšie mechanizmy konsenzu, aby sa zabránilo útokom Sybil, ktorý zahŕňa vytvorenie viacerých falošných identít v sieti typu peer-to-peer.

Na druhej strane, súkromné blockchajny sú založené na mechanizme kontroly prístupu. Umožňujú pripojenie len autentifikovaných alebo overených účastníkov. Teda existuje vlastník alebo správca siete, ktorý kontroluje pravidlá prístupu k blockchainu. Súkromné blockchajny sú v podstate uzavreté distribuované účtovné knihy fungujúce ako zabezpečené databázy.

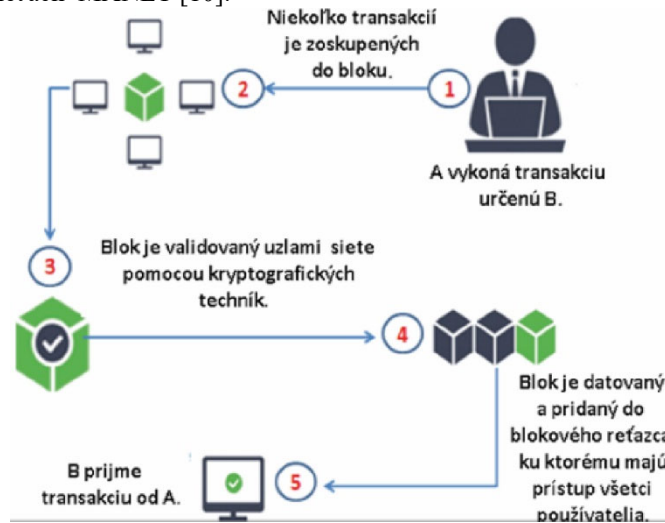
Je potrebné spomenúť tiež povolený blockchain ako tretí typ prístupu odlišný od súkromného a verejného. Povolenie sa môže v určitých súvislostiach vzťahovať na udeľovanie povolení účastníkom iba na konkrétne operácie, ako je čítanie alebo zápis do reťazca. Pokiaľ nie je uvedené inak, budeme odkazovať na povolené blockchajny v rovnakom význame ako na súkromné. Medzi týmito dvoma modelmi stojí hybridný model, ktorý kombinuje verejný aj súkromný Blockchain. Napríklad vybrané transakcie je možné odoslať zo súkromného do verejného Blockchainu pre otvorený prístup a bezpečný pôvod dát [12].

#### A. Blockchain a viacprekovaná mobilná ad-hoc sieť MANET

Jedným zo základných problémov pri smerovaní paketov v MANET je nedostatok dôvery a reputácie zúčastnených uzlov, čo často vedie k nespoľahlivému doručovaniu paketov. Tento model používa frakciu uzlov na overenie smerovacích akcií prijatých inými uzlami a distribuovaný konsenzuálny mechanizmus v sieťach Blockchain na získanie reputácie každého uzla v sieti [11].

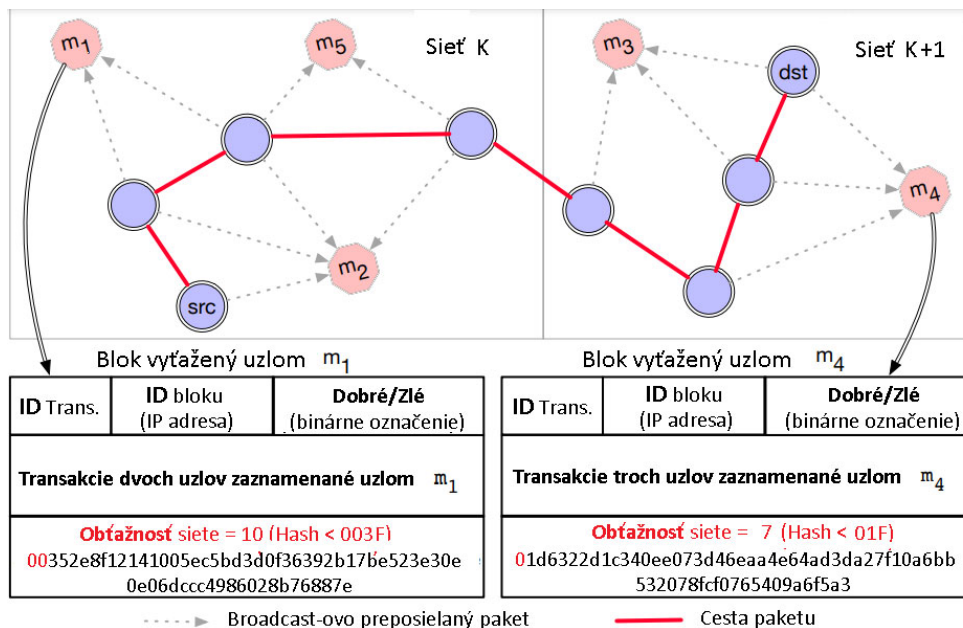
Konkrétne využíva heterogénne obtiažnosti pre dosiahnutie dôveryhodnosti overenia a navrhnutie bodovacieho systému na izoláciu škodlivých uzlov, a to prostredníctvom algoritmu Proof of Work. Reputácia uzla je založená na kombinácii úrovne obtiažnosti a skóre. Táto reputácia je zakomponovaná do novej metriky smerovania na výpočet najkratšej a najdôveryhodnejšej cesty medzi zdrojovým a cieľovým uzlom. Cieľom je odradiť škodlivé uzly vylúčením tých, ktoré sa zúčastňujú smerovania paketov.

Spoločná simulácia blockchainu a smerovacieho algoritmu odhalila  $\approx 12\%$  zlepšenie celkového doručenia paketov za prítomnosti smerovacích útokov v porovnaní s konvenčnými smerovacími algoritmi v sieťach MANET [10].



Obr. 3 Princíp činnosti Blockchainu v MANET [10].

Tento model je navrhnutý tak, aby priradil reputáciu uzlu na implementáciu spoľahlivého a dôveryhodného smerovania paketov. Konvenčné smerovacie protokoly v mobilných viacpreskokových sieťach sú založené na teórii najkratšej cesty medzi dvojicou zdroj a cieľ, ktorá je často reprezentovaná počtom preskokov. Cieľom je preformulovať tieto náklady na smerovanie s informáciami o reputácii uzlov v procese smerovania [2].



Obr. 4 Ukážka činnosti Blockchainu v MANET [2]

Každá MANET sieť teda predstavuje oblasť dosiahnuteľnosti ťažiarov z každého mobilného uzla v rámci tejto siete tak, že pakety poslané uzlami prijíma najmenej jeden ťažiar v okolí, na Obr. znázornené ako bodkované šípky. Každý ťažiar sleduje akcie vykonané uzlami pri smerovaní paketov, označuje ich ako dobré alebo zlé akcie a zaznamenáva ich ako transakcie. Tieto transakcie sú ťažiarom agregované a použité na vytvorenie bloku. V tejto práci je úloha Blockchainu dvojaká [11], [12], [13]:

- 1) poskytuje nemenný záznam o správaní uzlov,
- 2) stanovuje požiadavku, aby náročnosť ťažby bola úmerná počtu ťažiarov v sieti, pôsobí ako opatrenie dôveryhodnosti validácie.

Vyťažené bloky sú zdieľané medzi ťažiarimi, aby sa priblížili k bloku, ktorý sa ťaží s najvyššou

obťažnosťou, pretože sa považuje za naj dôveryhodnejší overený blok. Je to definované ako konsenzus MostDifficult-Chain. Akonáhle každý ťažiar získa aktualizovaný stav blockchainu, vysielá ho do uzlov v sieti, aby tiež aktualizovali svoje miestne stavy. Informácie uložené v blockchaine sa používajú na výpočet reputácie každého uzla, ktorý sa má použiť pri výpočte budúcich trás [2].

## I. ZÁVER

Technológia Blockchain si v poslednom čase získava veľa priestoru vo výskumných aktivitách. Vedci v oblasti MANET študujú a používajú Blockchain na riešenie problémov so zabezpečením a spoľahlivosťou prenosu. Blockchain je teda distribuovaná databáza udržiavaná viacerými uzlami v sieti a používa sa ako riešenie problémov dôvery a bezpečnosti. Teda hlavný dôraz výskumných aktivít v MANET sieti je zvýšenie dôveryhodnosti a reputácie jednotlivých mobilných uzlov, ktoré sú súčasťou danej siete. Práve technológia Blockchain sa ukazuje ako veľmi zaujímavý nástroj na dosiahnutie daných vedeckých cieľov. Výsledky mnohých výskumov v oblasti bezpečného smerovania založeného na technológii Blockchain ukazujú, že daná technológia pomáha pri detekcii akejkoľvek škodlivej aktivity v mobilných viacpreskokových sieťach.

## POĎAKOVANIE

Táto publikácia vznikla vďaka podpore Agentúry na podporu výskumu a vývoja v rámci projektu c. APVV-17-0208 - Odolné mobilné siete na doručovanie obsahu.

## REFERENCIE

- [1] C. E. Perkins, "Ad hoc networking: an introduction," in Ad hoc Networking, pp. 1–28, 2001.
- [2] M.A.A. Careem, A. Dutta, "Reputation based Routing in MANET using Blockchain,," International Conference on COMMunication Systems & NETworks (COMSNETS), 2020.
- [3] I. Woungang, S. K. Dhurandher, M. S. Obaidat GE, and R. D. Peddi, "A DSR-based routing protocol for mitigating blackhole attacks on mobile ad hoc networks," Security and Communication Networks, Vol. 9, No. 5, pp. 420–428, 2016.
- [4] A.K. Das, R. Chaki, K. N. Dey, "Secure energy efficient routing protocol for wireless sensor network,," Foundations of Computing and Decision Sciences, Vol. 41, No. 1, pp. 3–27, 2016.
- [5] B. Sen, M. Goldie Meitei, K. Sharma, M. Kanti Ghose, S. Sinha, "A trust-based intrusion detection system for mitigating blackhole attacks in MANET," in Advanced Computational and Communication Paradigms, pp. 765–775, 2018.
- [6] H. Moudni, M. Er-rouidi, H. Mouncif, and B. El Hadadi, "Performance analysis of AODV routing protocol in MANET under the influence of routing attacks," in Proceedings of the International Conference on Electrical and Information Technologies (ICEIT), 2016.
- [7] A. A. Hashim, M. M. Farhan, S. Alshybani, "Performance Evaluation of OLSR and AODV Routing Protocols over Mobile Ad-hoc Networks," 2019 First International Conference of Intelligent Computing and Engineering (ICOICE), pp. 1-8, 2019.
- [8] S.S. Anjum, R. Md Noor, M.H. Anisi, "Review on MANET based communication for search and rescue operations.," Wireless Personal Communications, Vol. 94, pp. 31–52, 2017.
- [9] A. Moussaoui, A. Boukream, "A survey of routing protocols based on link-stability in mobile ad hoc networks.," Journal of Network and Computer Applications, Vol. 47, pp. 1–10, 2017.
- [10] G. Liu, H. Dong, Z. Yan, X. Zhou and S. Shimizu, "B4SDC: A Blockchain System for Security Data Collection in MANETs," in IEEE Transactions on Big Data,
- [11] Q. Xing, B. Wang, W. Baosheng, "BGPcoin: Blockchain-Based Internet Number Resource Authority and BGP Security Solution," *Symmetry* Vol. 10, No. 9: 408.
- [12] V. L. Narayana, D. Midhunchakkaravarthy, "A Time Interval based Blockchain Model for Detection of Malicious Nodes in MANET Using Network Block Monitoring Node," Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020, pp. 852-857, 2020.
- [13] A. Jarjis, G. Kadir, "Blockchain Authentication for AODV Routing Protocol," Second International Conference on Blockchain Computing and Applications (BCCA), pp. 78-85, 2020.

# Livestreaming Delivery with Enhanced Interactive Functions in NTI in SANET

<sup>1</sup>František JAKAB, <sup>2</sup>Lucia MICHALENKOVA, <sup>3</sup>Dávid CYMBALÁK, <sup>4</sup>Roman VÁPENÍK

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

<sup>3,4</sup> Slovak Centre of Scientific and Technical Information, Slovak Republic

<sup>1</sup>frantisek.jakab@tuke.sk, <sup>2</sup>lucia.michalenkova@student.tuke.sk, <sup>3</sup>david.cymbalak@cvtisr.sk, <sup>4</sup>roman.vapenik@tuke.sk

**Abstract** — This paper deals with the interaction with the viewer while watching live streaming or on demand videos in specific environment of National Telepresence Infrastructure (NTI) livestreams environment. The aim of this work is to analyze existing solutions and based on the acquired knowledge to propose own solution in the form of a responsive web application. The part of the proposed solution was also an administration section in which it is possible to manage livestream, polls and chat data. For development of this application video and livestream data of NTI were involved and interconnected. NTI is robust videoconference and video streaming network running on Slovak academic data network ( SANET ). To verify the usability of the implemented system, the solution has been user-tested in academic environment.

**Keywords** — Interaction with user, Poll, Streaming, Voting, Real-time communication, Web application

## I. INTRODUCTION

Thanks to this technological advancement, which has made possible to use the website to provide video content and live stream, it has become possible to expand the audience of the content presented by participants who cannot physically participate but they are interested in the content. Although, such a viewer has the opportunity to receive the presented content in a high quality connection comparable to the physically present viewer, it is necessary to consider other possibilities of interaction with the audience such as discussion, opportunity to ask questions, provide feedback etc. These disadvantages for the viewer, who watches the presented content using Internet technologies, have gradually found their solutions in the form of additional functionalities implemented in web portals focused on video content sharing. The discussion is represented in a virtual form by the usual functionalities, such as commenting through comments and chat during live stream. Conversely, to express feedback, a number of different types of functionality arise, such as evaluation based on a numerical scale in the form of interactive stars element or the possibility of tagging with Like or Dislike, creating lists of favorite videos, or supporting the authors and other forms.

The goal of this work is to analyze the available and possible functionalities intended for interaction with the viewer during live streaming or video recording in a web environment. The result of this analysis should be the implementation of selected functionalities into its own web system, for which video data of the National Telepresence Infrastructure CVTISR were provided in the form of API.

## II. ANALYSIS OF EXISTING SOLUTIONS

Slovak National Telepresence Infrastructure is project, which associates videoconference for academic and research institutions in Slovakia based on SANET nodes. Web platform of National Telepresence Infrastructure consists of 3 interconnected web portals. NTI News, NTI Calendar and NTI Archive. The NTI Archive portal contains video on demand section of public recordings, live streaming section and section for administration of recordings and broadcasts with statistics [1]. Users have ability to comment and rate videos from NTI Archive. Based on the knowledge from the analysis of existing systems, functionalities that need to be implemented within the new

system have been defined. One of the main requirements for the new system is implementation of administration section for ability to manage content of system. Because of that, it is necessary to define the user roles, and based on that, it is possible to distinguish what rights the given user has and what functionalities are available for him. Only a user with administrator rights will be able to access the administration section, which contains a list of all videos by default. The following actions can be performed in this section: Add new video, remove existing video, hide existing video, edit the attributes of an existing video, access the list of user added comments for specific video and delete any of them, create a poll and access its results, lock the video content on specific password. Beside the listed actions, user with admin rights has access to all other sections to which other users of the system have access. Users of the system, no matter whether they are registered, will be able to access and perform the following actions: View live and archived videos from the video portal, view videos which are accessible only after entering a correct password, rate videos, view comments, view polls and vote, view live chat, share video via social networks. Registered and authenticated users will be able, in addition to the above-mentioned actions also: add comments. add messages to chat, add videos to 'your favorites list' and access this list and remove videos from their list. In order to provide user not only with content that he selected, but also other content of the system that could be interesting and relevant for him, recommended sections have been included into system design and implementation.

### III. DESIGN AND IMPLEMENTATION OF SYSTEM

The architecture of the developed system is based on server and client application interconnected with database and video sources from NTI API. Users of the system, both with and without admin rights, access the client application, which is available as responsive web application. Client application communicate with the server part using the secure HTTP protocol. Server application communicate with the database and provide those data to the client application.

For implementing client application, JavaScript framework Vue.js have been chosen. Before starting with implementation, following actions have been made: Object and Actions Analysis, Conceptual model creation, Screen Sequence Diagram creation. Vue.js has been inspired by Model-view-viewmodel architectural pattern, but as convention is used the variable vm (short for ViewModel) to refer to Vue instance. Every Vue application starts by creating a new Vue instance. The architecture of the Vue.js application framework works on the principle of creating applications from small, separate and often reusable components, which are Vue instances. During implementation of designed client application 28 Vue components were created and 2 Mixin classes, in addition for implementations methods that need to be accessed in various components.

For server side, Node.js framework Express.js with Mongoose library have been used for implementation. Server application consists of definition of database models, controllers classes and routers. HTTP requests from client side are routed by application using defined routers to the appropriate controller. Controllers contain implementation of methods that specify actions that need to be done within internal database and return data to client side.

For storing and working with data, the database management system MongoDB has been chosen. MongoDB belongs among non-relational databases called NoSQL and uses JSON-like documents with schema. Figure 1 presents all schema created for implemented application.

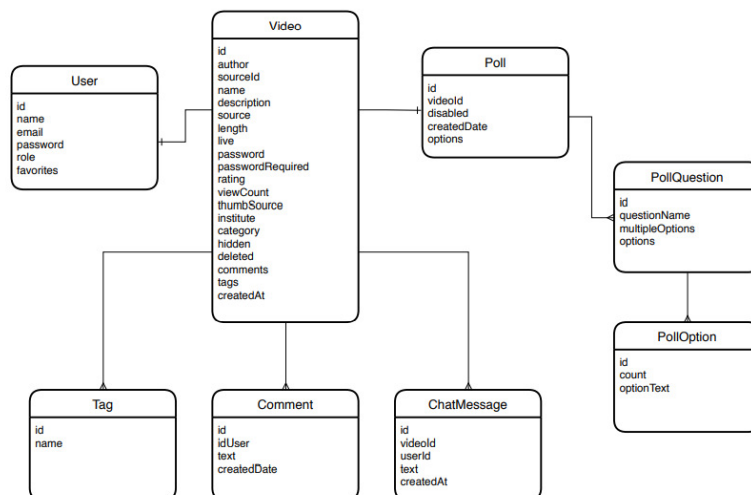


Fig. 1 Database schema of designed system

Communication between client and server is done using secure version of HTTP protocol. From client side, JavaScript library Axios has been used for implementation of XMLHttpRequests, which allows retrieve data without having to do a full-page refresh, which enables to update just part of a page without disrupting what the user is doing.

For implementing interactive functions as chat or poll, it is important to use real time communication. In 2011 Websocket protocol has been standardized, which as Furukawa [2] mentioned, allows asynchronous full-duplex communication over TCP connection.

Socket.io library has been used for implementing Websocket communication, which provides following advantages: Connections are established even in the presence of proxies and load balancers, personal firewall and antivirus software. Unless instructed otherwise, a disconnected client will try to reconnect forever, until the server is available again. As Walsh presented [3], socket.io provides fallback mechanisms, such as Flash Socket, AJAX long-polling, AJAX multipart streaming, IFrame and JSONP polling, in case Websocket communication cannot be established. Socket.io supports broadcast, which allows to send a message to all connected clients, without the need to create a list of all connected clients, as it is in the case of Websocket. Support of specific channels, called Rooms, that sockets can join and leave. This allows broadcast to any given room, reaching every socket that has joined it. The designed application may contain several different chats or polls, so the mentioned Rooms will be used to divide communication and broadcast the messages only to those clients, where it is needed. The schema of such implementation is presented in Figure 2.

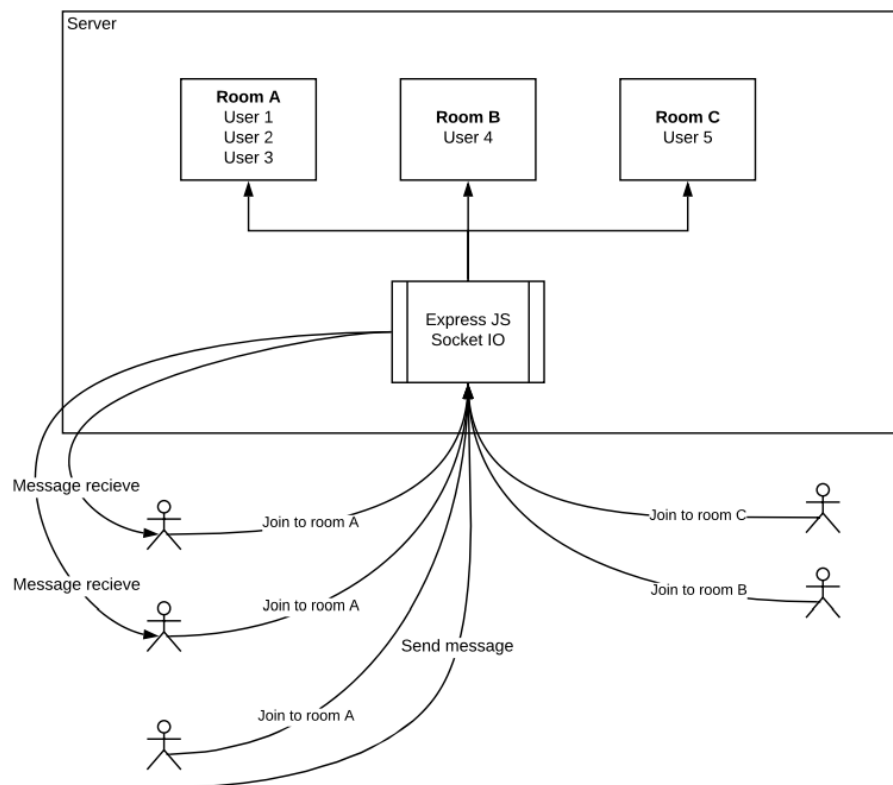


Fig. 2 Schema of implementation real-time communication using socket.io library.

In order to verify the functionality and usability of the implemented solution, a user testing was performed for both parts of the system.

Usability Testing of user sections was performed with 20 participants, for whom were given a brief description of the application and simple scenarios aimed at testing the most important functionalities of the system. These scenarios were defined very generally without detailed steps to verify that the user interface is intuitive, and the user can easily navigate through system, without the need for additional explanation. After testing, participants were asked to fill in a questionnaire. The first part of questionnaire was focused on the evaluation of defined scenarios. For the second part of questionnaire, the methodology of SUS - System Usability Scale [4] was used for evaluating the usability of the application. The SUS score calculated based on the provided answers was 89.2, that is above the average score, which is according the 500 studies [5] 68. User interface for this section is depicted on Figure 3 and Figure 4



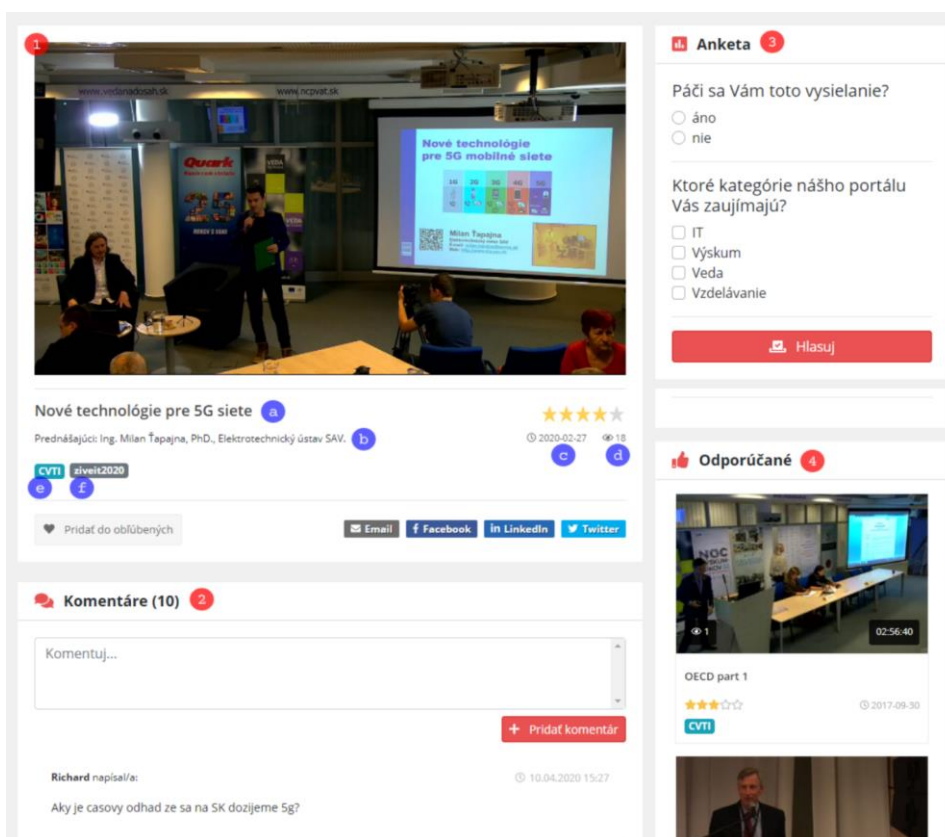


Fig. 3 Implementation of polls functionality in livestream environment.

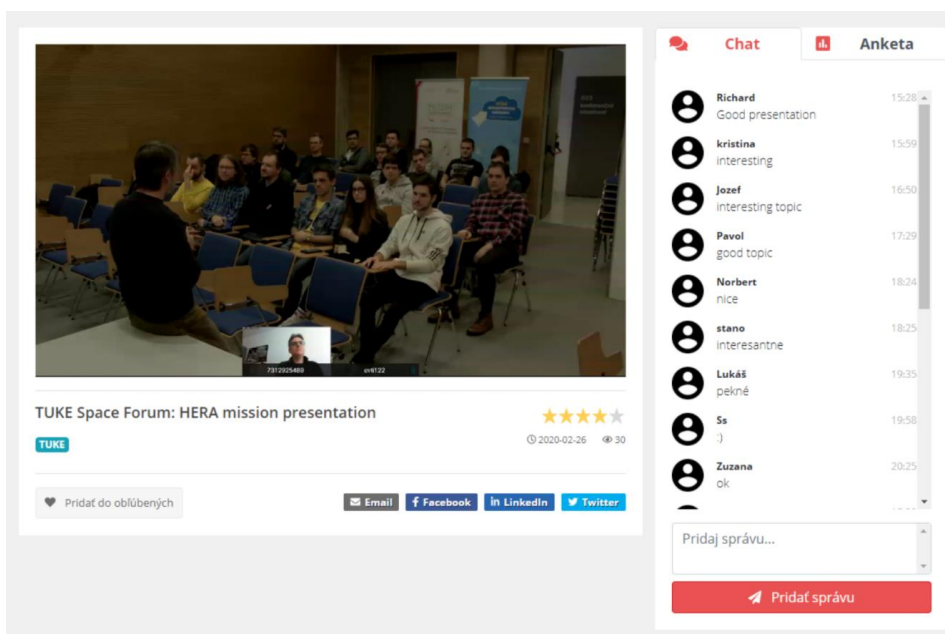


Fig. 4 Implementation of experimental chat functionality in livestream environment.

Admin section testing was performed with one user who received the credentials of a user with admin rights and defined tasks. User was asked to perform these tasks without explanation how the system works and where and how each action can be performed. Based on the feedback provided by the user after completing the tasks, it was found that the defined tasks could be performed successfully, but the user provided feedback for improvement of some areas. Part of user interface of admin section is depicted on Figure 5.

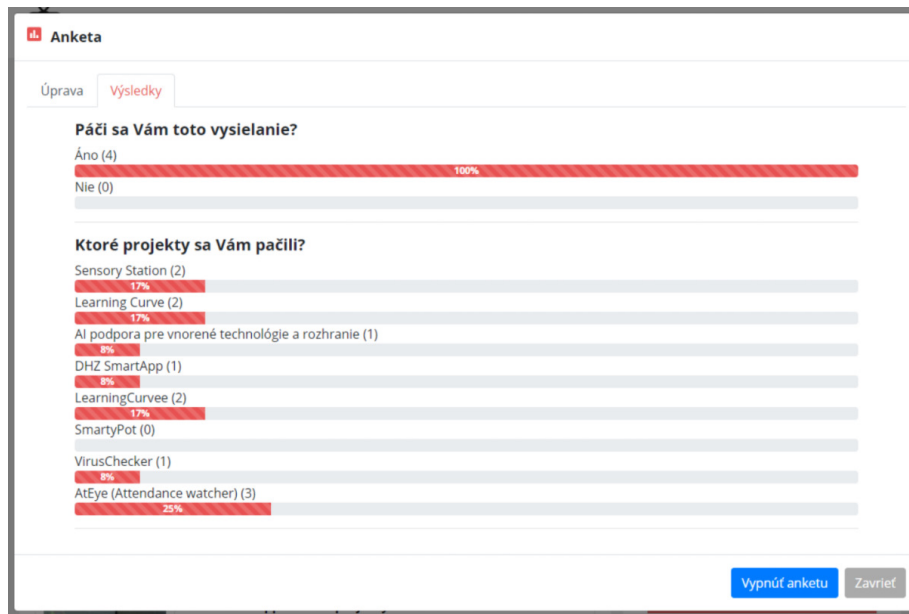


Fig. 5 Polls results interface in admin section of designed solution.

#### IV. CONCLUSION

The result of the work is the web-based video system, consisting of two separate applications, using Javascript programming languages and real-time communication protocol. The purpose of the system is to provide live streams listeners with enhanced interactive functions, such as chatting during live stream, commenting on videos from archive, voting in polls, rating the videos, adding videos to list of favourite videos and sharing via email or social networks as Facebook, LinkedIn and Twitter. The important part of the system is administration sections, which gives ability to add new and update existing content, manage the existing comments and create polls and observe their in real time updated results.

The both parts of the system have been user tested and provided functionality and usability of set objectives with future possibility to enhance current system with additional features.

#### ACKNOWLEDGMENT

This publication has been published with the support of the Operational Program Integrated Infrastructure within project: Research in the SANET network and possibilities of its further use and development, ITMS code: NFP313010W988, co-financed by the ERDF.

#### REFERENCES

- [1] D. Cymbalak, F. Jakab, M. Michalko, R. Vapenik, J. Turna, L. Bilsky, and J. Kovacova, "Effective collaborative environment with live broadcasts using slovak national telepresence infrastructure," pp. 1–6, 10 2017.
- [2] Y. Furukawa, "WEB-BASED CONTROL APPLICATION USING WEBSOCKET," 2011. [Online]. Available: <http://epaper.kek.jp/icalpecs2011/papers/wemau010.pdf>
- [3] D. Walsh, "WebSocket and Socket.IO," 2010. [Online]. Available: <https://davidwalsh.name/websocket>
- [4] J. Brooke, "SUS - A quick and dirty usability scale," Usability evaluation in industry, pp. 189–194, 1996.
- [5] J. Sauro, "Measuring usability with the system usability scale (SUS)," 2011. [Online]. Available: <https://measuringu.com/sus/>

# Multikamerová distribúcia streamovaného videa z UAV pri sledovaní pohybujúceho sa objektu

<sup>1</sup>František JAKAB, <sup>2</sup>Dávid CYMBALÁK, <sup>3</sup>Miroslav MICHALKO, <sup>4</sup>Martin KORÉNEK

<sup>1,2</sup> Univerzitný vedecký park Technicom, Technická univerzita v Košiciach

<sup>3</sup> Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

<sup>4</sup> Centrum vedecko-technických informácií SR

<sup>1</sup>frantisek.jakab@tuke.sk, <sup>2</sup>david.cymbalak@uvptechnicom.sk,

<sup>3</sup>miroslav.michalko@tuke.sk, <sup>4</sup>martin.korenek@cvtisr.sk

**Abstrakt** — Článok sa zaoberá využitím modelu automatický výberu najviac relevantného záberu na pohybujúci sa objekt v multikamerovom streamingovom toku generovanom kamerami bezpilotných lietadiel ( UAV ). Na základe analýzy adaptívnych streamingových protokolov s prepojením na mechanizmy počítačového videnia pri sledovaní polohy objektu v obraze bola navrhnutá adaptácia hodnotiacej metriky pre jednotlivé streamingové toky obrazu z UAV. Evaluácia metriky je založená na základe polohových a kompozičných parametrov vzhľadom na prekryv záberov v UAV pohybujúcej sa sústave. Mechanizmus evaluácie záberov zabezpečí indikáciu najviac relevantného záberu na objekt a okamžitú komunikáciu so streamingovým serverom pre zmenu výstupného streamingového toku sústavy UAV za aktuálne najviac hodnotený. Popisovaný model bol experimentálne overený v náhradných podmienkach.

**Kľúčové slová** — UAV, Streaming, Dron, Detekcia objektov vo videu

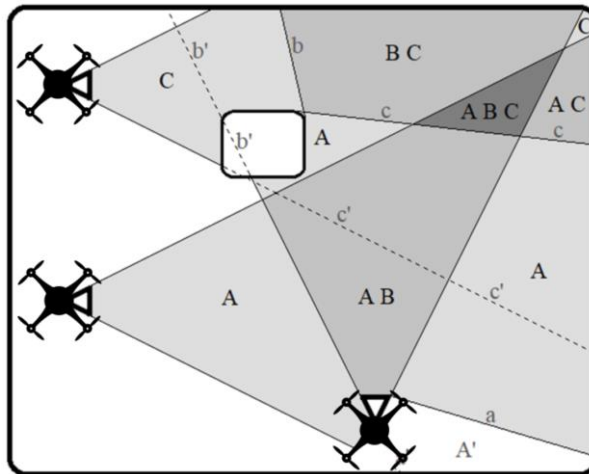
## I. ÚVOD

Sledovanie objektu v živom videu z priestoru pokrytom sieťou kamier je často skúmaným odvetvím najmä bezpečnostných systémov. Primárne je cieľom týchto výskumov riešiť problémy vznikajúce pri detekcii a trasovaní objektu v rôznorodých záberoch rôznej farebnosti a z rôznych uhlov. Pri detekcii objektu pomocou kamier umiestnených na bezpilotných lietadlách ( UAV ) sa otvára široké spektrum vznikajúcich problémových oblastí a ich riešenia. Cieľom tohto výskumu je adaptácia automatického modelu, ktorý na základe periodického ohodnocovania všetkých záberov streamovaných z kamier viacerých UAV, vyhodnocoval distribuované toky a na výstupe indikoval vždy iba najviac relevantný záber na objekt z vybraného UAV v reálnom čase. Je potrebné navrhnuť a prispôbiť hodnotiacu metriku streamovaných zdrojov, tak aby bolo hodnotenie v reálnom čase závislé na aktuálnej polohe sledovaného objektu na záber, veľkosti objektu v zábere a na ďalších vybraných parametroch záberu. Samotnej adaptácii predchádzala analýza súčasných streamingových technológií pre účely konfigurácie mechanizmu riadenia streamingu a vhodných detekčných a sledovacích algoritmov pre účely zostavenia mechanizmu hodnotenia zdrojov [1] obrazových tokov z UAV.

## II. ANALÝZA PROBLEMATIKY MULTIKAMEROVÉHO PRENOSU OBRAZU V SÚSTAVE UAV

Rozsiahly priestor s pohybujúcim sa objektom je v reálnych podmienkach väčšinou obťažné sledovať jedinou kamerou z jediného UAV. Riešenia sa ponúka vo forme samotného riadenia pohybu UAV v závislosti od umiestnenia projektu a navádzania UAV po trajektórii v smere za objektom. Avšak toto riešenie poskytne kamerový pohľad na sledovaný objekt vždy iba z jedného uhla záberu a v jednej veľkosti objektu vzhľadom na záber okolia. Taktiež neeliminuje prekážky v zábere objektu a pri manévroch ich obchádzania sa sledovaný objekt stráca z dohľadu a dochádza k chybovosti. Vhodnejším spôsobom na sledovanie sa pohybujúceho objektu v určenom priestore je sledovanie pomocou sústavy viacerých UAV tvoriacich multikamerový

system, ktoré ponúka väčšiu škálovateľnosť a pokrytie priestorov v rôznom tvare a s rôznymi priestorovými obmedzeniami. Pri takomto multikamerovom systéme [2] sa zobrazenie objektu z pohľadu vybraného UAV môže výrazne líšiť od zobrazenia objektu z pohľadu ostatných UAV a to najmä z dôvodu rozdielného osvetlenia, otočenia alebo aj z dôvodu rozdielov samotných vlastností jednotlivých kamier. Pri prekrývajúcich sa záberoch je možné objekt sledovať pomocou komunikácie a odovzdávania informácií medzi jednotlivými uzlami multikamerového UAV systému (Obr. 1).



Obr. 1 Prekrýv záberov kamier pri multikamerovom spracovaní obrazu z UAV a detekcii objektov v zábere

Pri zónovom pokrytí a snímania priestoru z UAV s neprekrývajúcimi sa zábermi jednotlivých kamier je sledovanie objektu problematické hlavne z dôvodu, že pozorovania objektu sú často oddelené nielen v priestore, ale aj v čase. Na rozdiel od sledovacích prístupov pri použití jednej kamery sa v tomto prípade nemôže blízkosť v priestore a čase využiť ako relevantná informácia medzi jednotlivými uzlami systému [3]. Napríklad objekt odchádzajúci zo záberu prvej kamery sa nedostane ihneď na záber nasledujúcej kamery, ale je tam určitá časová aj priestorová odchýlka. Taktiež sledovaný objekt sa môže dostať zo záberu jednej kamery na záber nasledujúcej kamery rôznymi cestami, čo môže ovplyvniť to, z ktorej strany záberu sa objekt objaví v nasledujúcej kamere.

Pri voľbe vhodných streamingových technológií je nutné vybrať vhodný video štandard / kodek, ktorý bude využitý prostredníctvom enkódera, ktorý v reálnom čase zabalí obrazový tok do kontajnera a odošle na streamingový server [4]. Dôležité sú ale najmä streamingové protokoly [5] a ich adaptácia na využitie v prenose obsahu z kamier umiestnených na UAV. Streamingové protokoly zabezpečujú v tomto prípade prenos paketov so živým multimediálnym obsahom [6] a podstatná vlastnosť pre účely využitia v tomto výskume je ich oneskorenie prenosu ( Obr. 2 )

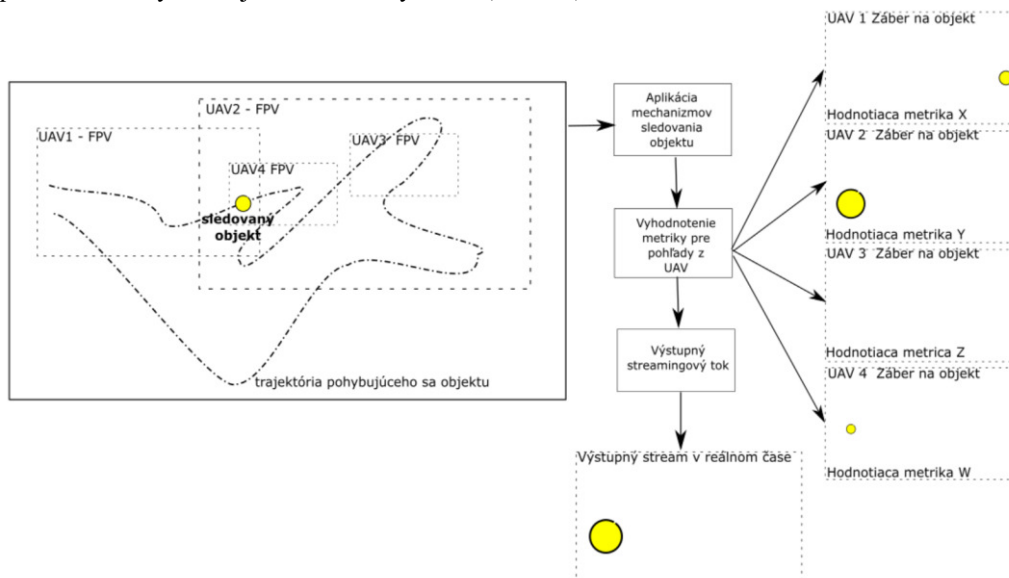
veľkosť oneskorenia	bežné cez web	redukované oneskorenie	nízke oneskorenie	blízke reálnemu času
typ vysielania	jednosmerné vysielanie veľkým skupinám	vysielanie správ či športových podujatí	streamovanie hier, e-športy	videokonferenčné hovory, ovládanie zariadení (PTZ kamery, drony)
čas. rozsah	45-18 sek.	18-5 sek.	5-1 sek.	menej ako 1 sek.
streamingové protokoly	HLS MPEG-DASH	vylepšené verzie HLS a DASH, RTMP	CMAF DASH, RTSP/RTP, „low latency“ HLS	WebRTC

Obr. 2 Analýza oneskorenia pri použití streamingových protokol pre účely prenosu obrazu v UAV [7]

### III. NÁVRH MODELU A EXPERIMENTÁLNE PROSTREDIE

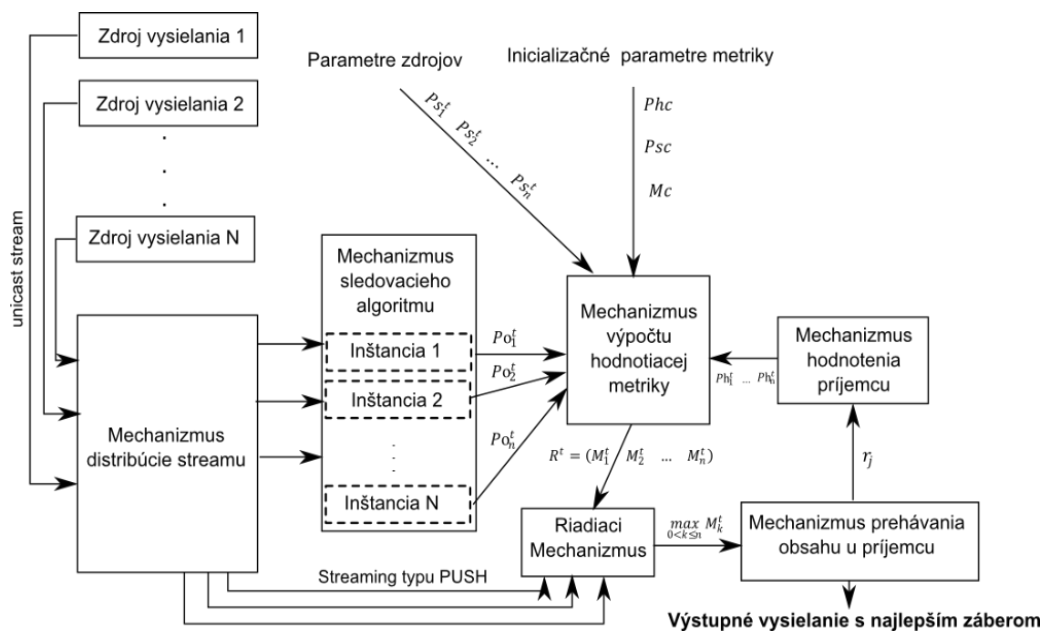
Štandardné prístup k detekcii pohybujúceho sa objektu v multikamerových sledovacích systémoch je formou indikácie objektu grafickým indikačným rámcom naprieč výstupmi zo všetkých kamier alebo zobrazia polohu objektu v globálnom pohľade na mape monitorovaného priestoru. Personál je teda nútený sledovať súčasne výstup zo všetkých kamier a hľadať v ňom zvýraznenú indikáciu alebo body detegovaného sledovaného objektu, respektíve sleduje súradnice a polohu na globálnej mape, v ktorej objekt reprezentuje iba určitý grafický bod a nie celé video. Z tohto dôvodu je pri prístupe nasadenia v kamerovom systéme viacerých UAV užitočným riešením prenášať snímaný obraz z jednotlivých UAV pomocou streamingových technológií a na výstupe zobrazit' v reálnom čase najviac relevantné živé video z UAV, ktoré má na sledovaný objekt najlepší uhol záberu a kompozíciu.

Pri návrhu adaptácie modelu vychádzajúceho z nášho predošlého výskumu multikamerového systému pre živé vysielania [1] sa vychádzalo z predpokladu, že jednotlivé zdroje živého vysielania z UAV z rôznych uhlov sú prenášané cez mechanizmus distribúcie streamov v podobe streamingového servera. Z distribučného mechanizmu sa prenesený tok každého záberu UAV videa dostáva k spracovaniu v mechanizme sledovacieho algoritmu v podobe jednej inštancie na každý zdroj videa. Každá inštancia mechanizmu pre sledovanie objektu záujmu vo videu posielala informácie o výsledku trasovania v reálnom čase a teda posielala vektor polohových a dimenzionálnych vlastností. Mechanizmus výpočtu hodnotenia taktiež dostáva na vstupe inicializačné parametre metriky obsahujúceho globálne konštanty jednotlivých zložiek metriky. Súčasne do mechanizmu výpočtu putuje vektor parametrov scény a kvality streamu odvíjajúci sa od použitia kódovania v mechanizme distribúcie alebo vstupnom enkodéry zdroja videa. Samotný mechanizmus výpočtu, ktorý priebežne zbiera vstupné parametre odosiela na výstup výsledkový vektor hodnotení všetkých vstupných zdrojov priamo do riadiaceho mechanizmu. Riadiaci mechanizmus v reálnom čase určí maximum hodnotenia a to, aký záber z konkrétneho UAV bude nastavený aktuálne na výstup vysielania. Riadenie je teda napojené priamo na mechanizmus prehrávania obsahu, kde sa z mechanizmu distribúcie zobrazí u príjemcu v prehrávači vždy len najviac relevantný záber ( Obr. 3 ).



Obr. 3 Návrh modelu multikamerového spracovania streamu z UAV (N=4)

Pri implementačnom návrhu bol na vstupe použitý živý tok obrazu zachytávaný N zdrojmi vysielania z UAV. Obraz z každého zdroja sa cez mechanizmus distribúcie streamu dostával prostredníctvom RTPS enkodéra priamo na WSE streamingový server. Následné sa každý distribuovaný RTPS stream prostredníctvom FFmpeg spracoval v separátnych inštanciách pyCMT v mechanizme sledovacieho algoritmu. Jednotlivé pyCMT inštancie odovzdávali výsledky o polohe a veľkosti sledovaného objektu systému pre výpočet finálnych metrik spolu s parametrami streamu z WSE servera a s parametrami hodnotenia príjemcov. Výsledné hodnoty metrik sa odovzdávali s informáciou maxima a identifikátora najvyššie hodnoteného streamu mechanizmu s vMix API. Odtiaľ sa následne posunul výsledný stream na výstup vo virtuálnom mixe, ktorý z WSE servera dostával každý tok separátne. Následne sa výstupný stream enkódoval na požadovaný formát mechanizmom prehrávania obsahu u príjemcu a dosiahol výstupné vysielania s najlepším záberom zo sústavy UAV ( Obr. 4 ).



Obr. 4 Štruktúra komunikácie a mechanizmov multikamerového spracovania streamu z UAV pri N uzloch

V simulačnom prostredí tvorenom prvotne náhradným riešením RTPS obrazových tokov boli realizované viaceré experimenty, pri ktorých sa snímal vybraný pohybujúci sa objekt záujmu viacerými pohybujúcimi sa kamerami súčasne v rôznych prostrediach s rôznou dynamikou záberu. Nasadený systém podľa navrhnutého modelu opätovne každý jeden zdroj video signálu duplikoval tak, aby jeden tok smeroval do mechanizmu pre výpočet metriky na báze sledovacieho algoritmu a druhý tok do streamingového mixu s enkodérom pre tvorbu výstupu vysielania. Výstupná obrazová metrika vypočítaná jednotlivo pre každý zdroj videa sa hodnotiacim mechanizmom získaním finálnych hodnotení pre všetky zdroje sa riadiacim mechanizmom našlo maximum, ktoré určilo, ktorý zdroj video signálu bude v danom momente na výstupe vysielania. Pri simuláciách predchádzajúcich implementácii na fyzických UAV sa v pilotnom riešení využije pripravené prostredia Gazebo a ROS (Obr. 5), kde sa virtualizuje kamerový výstup pri pohybe UAV vo virtuálnom prostredí.



Obr. 5 Použitie simulačného prostredia Gazebo a ROS

#### IV. ZÁVER

Výsledkom tejto publikácie je návrh adaptácie modelu automatickej distribúcie najviac relevantného záberu na objekt v multikamerovom systéme pozostávajúcom zo sústavy UAV. Adaptívna evaluačná metrika bola prispôbená formou použitia voľby váh jednotlivých zložiek metriky s prioritizáciou lokalizácie a veľkosti objektu. Parametre prenosu video a audio stopy boli v tejto adaptácii modelu segregované. Evaluačná metrika sa stala základom pre vytvorenie mechanizmu indikácie najviac relevantného záberu na pohybujúci sa objekt v sústave UAV s komunikáciou smerom na streamingový server. Experimentálna implementácia prebehla iba v náhradných podmienkach substitúcia UAV za kamerové pohybujúce sa moduly v streamingovom reťazci. Simulačné prostredie pre model je v príprave pre simulácie v prostredí Gazebo a ROS s budúcim transferom mechanizmu na reálne zariadenia.

Riešenie má potenciál prispôsobiť sa rôznym typom použitia ako napríklad UAV monitoring osôb a zvery pomocou sektorového snímania územia, monitoring dopravných prostriedkov pomocou siete UAV, interiérový UAV monitoring pohybu výrobných artiklov priemyselnej výroby v rozsiahlych skladoch a pod.

#### POĎAKOVANIE

Táto publikácia vznikla vďaka podpore v rámci Operačného programu Integrovaná infraštruktúra pre projekt: Inteligentné operačné a spracovateľské systémy pre UAV, kód ITMS: 313011V422, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

#### LITERATÚRA

- [1] CYMBALÁK D., VÁPENÍK R. a kol.: Multi-camera broadcasting model with automation of optimal scene switching. In: Infocommunications Journal. Vol. 8, no. 4, 2016 s. 15-22. - ISSN 2061-2079
- [2] VÁPENÍK R.: Non-Standard Situations in Multi-Camera Environment. In: SCYR 2017. - Košice : TU, 2017 s. 174-175. - ISBN 978-80-553-3162-1
- [3] VÁPENÍK R.: Object tracking in videoconferencing systems. In: SCYR 2015. – Košice : TU, 2015 s. 261-265. – ISBN 978-80-553-2130-1
- [4] JAKAB A., VÁPENÍK R.: Using modern streaming technologies for high quality and performance-oriented live streaming. In: SCYR 2014. – Košice : TU, 2014 s. 279-282. – ISBN 978-80-553-1714-4
- [5] JAKAB A., VÁPENÍK R.: High efficient live streaming using the latest available streaming technologies. In: Poster 2014. – Praha : ČVUT, 2014 s. 1-4. – ISBN 978-80-01-05499-4
- [6] CYMBALÁK D., VÁPENÍK R. a kol.: Enhanced Interoperability with New Collaborative and Streaming Functions Implemented in Slovak National Telepresence Infrastructure (NTI). In: 16th International Conference on Emerging eLearning Technologies and Applications (ICETA). – Starý Smokovec/Košice : IEEE, 2018 s. 105-110. - ISBN 978-1-5386-7914-2
- [7] WOWZA: Video Streaming Latency Report. Dostupné online. <https://www.wowza.com/blog/2019-video-streaming-latency-report>

# Simulation of Activities in a Robotic Swarm

<sup>1</sup>Patrik SCHÜTZ, <sup>2</sup>Ján VAŠČÁK

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

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

<sup>1</sup>patrik.schutz@student.tuke.sk, <sup>2</sup>jan.vascak@tuke.sk

**Abstract** — The aim of this paper is to describe the search of the optimal path using robotic swarm in simulated environment with obstacles. The paper deals with the implementation of Particle Swarm Optimization and Ant Colony Optimization methods and their modifications. The paper also describes the implementation of potential fields and diffusion model. The functionality of the implementations is verified on a series of performed experiments.

**Keywords** — Ant colony optimization, particle swarm optimization, path planning, robotic swarm

## I. INTRODUCTION

The world of the present and the world of the future are increasingly shifting towards robotics and artificial intelligence. Robots are slowly but surely becoming part of everyday life at home, at work or in services. However, robots cannot plan a path when moving in space, as humans, animals, insects, or other living beings do. By observing insect, animal or human communities, it is possible to simulate the behavior of these communities for various purposes and issues, as well as path planning issues. The purpose of this paper is to find the optimal path in a simulated environment with obstacles with the help of a robotic swarm and swarm intelligence approaches.

A two-dimensional simulation environment was designed, created, and simulated, while maps were created of the type metric map using the occupancy grid method. Experiments of Particle Swarm Optimization and Ant Colony Optimization methods were performed on an environment map with grids and densely placed obstacles.

## II. PARTICLE SWARM OPTIMIZATION

An initial population with  $n$  numbers of particles is initialized. Each particle is assigned a starting position that is the same as the starting grid, or the starting position of one search particle is assigned, and the other particles are assigned a random starting position ranging from the upper and lower bounds of the map width for coordinate  $x$  and the upper and lower bounds of the map length for coordinate  $y$  so that the particle is not at the obstacle position. Each particle is assigned an initial velocity for both coordinates in the range  $(-1, 1)$ . Initialization of the population is followed by a main loop. In the loop, the position of each particle in the searched area is initially evaluated using the fitness function. After calculating the fitness of the position of each particle in the swarm, the best position within the swarm is compared with the current positions of all particles. If the current position of a swarm member is better, it will be replaced by this position. The calculation of the new velocity in PSO for a particle is performed using the equation [1]:

$$V_i(t+1) = w * V_i(t) + c_1 * r_1 * (p_{best} - x_i(t)) + c_2 * r_2 * (g_{best} - X_i(t)), \quad (1)$$

where  $w$  is the inertial weight,  $V_i(t)$  is the current velocity,  $c_1$  is the cognitive coefficient,  $c_2$  is the social coefficient,  $r_1$  and  $r_2$  are random values from the range  $(0, 1)$  and  $X_i(t)$  is the current position. Based on the calculation of the new particle velocity, the new particle position is calculated according to the equation:



$$X_i(t + 1) = X_i(t) + V_i(t + 1) \quad (2)$$

The grid that corresponds to its new position is then selected to move the particle. Otherwise, the grid closest to the newly calculated position is selected. After updating the new positions, the loop repeats until the maximum number of particle moves is reached or a solution is found. In each iteration, the particle can move by one grid.

#### A. Time Varying Acceleration Coefficients and Time Varying Inertia Weight

The acceleration coefficients  $c_1$ ,  $c_2$  and the inertia weight  $w$  are constants in the basic PSO algorithm. To improve the performance of PSO, a modification is used where the acceleration coefficients and the inertia weight are time-varying (TVAC-IW). The TVAC-IW has the same basis as the PSO, except that when calculating the new particle velocity, the new acceleration coefficients and the new inertia weight are calculated first. The upper and lower limits of the values for the acceleration coefficients are set. The cognitive coefficient  $c_1$  is initially set-up to the upper limit and the social coefficient  $c_2$  is initialized to the lower limit. With each iteration, the value of the cognitive coefficient decreases and the value of the social coefficient increases so that the final value of the cognitive coefficient is the lower limit and the final value of the social coefficient is the upper limit. In each iteration, new acceleration coefficients are calculated. The inertia weight is initialized to the upper limit, with the value of the inertia weight decreasing with each iteration, up to the lower limit [2].

#### B. Standard PSO 2011

Modification Standard PSO 2011 (SPSO2011) uses rotational invariance. The difference between the PSO and SPSO2011 algorithm is in the calculation of new velocities for particles, where for each particle in each iteration the center of gravity is calculated around three points: the current position of the particle, a point just behind the best personal position of the particle and a point just behind the best position within the swarm. The center of gravity is used to calculate a random position in the hyper-sphere [3].

#### C. Quantum Delta-Potential-Well-based PSO

Modification Quantum Delta-Potential-Well-based PSO (QDPSO) is an algorithm using a probability search technique, or in other words, it is an uncertainty search algorithm. Based on the characteristics of the Delta-Potential-Well wave function, some particles occur at a greater distance from the local attractor during searching, which is forbidden in the basic PSO. However, these particles are more likely to find a better position when searching. QDPSO is based on the idea that particles have only a position vector and no velocity vector. In the case of QDPSO, therefore, the speed is not calculated, which reduces the number of required parameters to one [4].

#### D. Barebones PSO

Studies have shown that particles in PSO converge to the weighted position of their best position and best position within the swarm. The modification of Barebones PSO (BPSO) is based on this convergence characteristic. The particles in BPSO do not have a velocity vector, they only have a position vector, so the acceleration parameters  $c_1$ ,  $c_2$  and the inertia weight  $w$  are eliminated. The new position of the particle is calculated using the Gaussian probability distribution [5].

### III. ANT COLONY OPTIMIZATION

An initial population with  $m$  numbers of ants is initialized. Each ant is assigned a starting position that is the same as the starting grid. Initial pheromone values are initialized for each grid in the searched space. Initialization of the population and pheromones is followed by a main loop. In the loop, each ant searches for a path to the target destination, and the probability of selecting an adjacent grid  $j$  to move from the current grid  $i$  is calculated based on the equation [6]:

$$p_{ij}^k(t) = \frac{[\tau_{ij}(t)]^\alpha * [\eta_{ij}(t)]^\beta}{\sum_{l \in N_i^k} [\tau_{il}(t)]^\alpha * [\eta_{il}(t)]^\beta}, \quad \forall j \in N_i^k, \quad (3)$$

where  $\tau_{ij}(t)$  represents the value of the pheromone on a grid,  $\eta_{ij}(t)$  represents the heuristic value,  $\alpha$  is the pheromone factor, which represents the relative importance of the pheromone,  $\beta$  is the heuristic factor, which represents the relative importance heuristics and  $N_i^k$  represents the available adjacent grids for moving for the  $k$ -th ant.

In the case of the occupancy grid method, it will always be  $\eta_{ij}(t) = 1$ , because the distance to move from the current grid to each adjacent grid is for all grids the same. The ant moves randomly to the adjacent grid according to the probability, so the higher the probability of the grid, the greater the chance of choosing this grid is. However, the ant cannot move to the grid of its previous position or to the grid occupied by an obstacle.

The ant search the area and selects the grids to move until it finds the target grid or the maximum number of possible moves of the ant is reached. Once all the ants in the colony have finished their search, the ants that have found a solution will place the pheromones on the grids they traveled on to find a solution. This process is called global updating of pheromones:

$$\tau_{ij}(t+1) = (1 - \rho) * \tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ij}^k(t), \quad (4)$$

where  $\rho$  represents the evaporation coefficient of the pheromone,  $\tau_{ij}(t)$  represents the current value of the amount of pheromone on the grid  $ij$ ,  $m$  represents the number of ants,  $\Delta\tau_{ij}^k$  represents the deposition of the pheromone  $k$ -th ant on the grid  $ij$ :

$$\Delta\tau_{ij}^k(t) = \begin{cases} \frac{Q}{L_k} & \text{if ant } k \text{ travels on tile } ij \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where  $Q$  represents the quantity of the stored pheromone and  $L_k$  represents the path length of the  $k$ -th ant that found the solution. After storing the pheromones, the values of the pheromones on the grids change, as do the probabilities of selecting a grid to move. The grid through which most ants passed when finding a solution has a higher value of pheromones. The grid through which fewer ants passed when finding a solution has a lower value, and the grid through which no ant passed has the unchanged value of pheromones. Increasing the amount of pheromone on the grid also increases the probability of its selection. Pheromone evaporation prevents endless accumulation of the pheromone.

#### A. Elitist Ant System

The ants in the Elitist Ant System (EAS) place the same amount of pheromones on the grids, as in the basic Ant System method. The difference is that the ant that finds the shortest solution in all iterations (the elite ant) puts twice the pheromone in its path. Whenever an ant finds the best solution in all iterations, the best solution gets a larger amount of pheromone, making the path of the best solution longer relevant, while worse solutions evaporate faster. The EAS method can accelerate convergence to an optimal solution, but it can also lead to finding and stagnating a sub-optimal solution. It is important to note that when finding a sub-optimal solution and then stagnating on it, there is also the possibility of detaching from this sub-optimal solution [7].

#### B. Ranked-based Ant System

The Ranked-based Ant System (ASrank) ranks all found ant solutions from best to worst, and the amount of pheromones is stored according to their path length. On the path of the best solution (on the shortest path), the ant stores the largest amount of pheromones, on worse paths (on longer paths), the ant stores a smaller amount of pheromones and on the worst path (on the longest path), the ant does not store any amount of pheromones. ASrank seeks optimal solutions efficiently by making a difference in the amount of pheromone stored between the solutions, with the solutions evaporating gradually. The worst solutions evaporate the fastest (the longest routes), and the best solutions evaporate the slowest (the shortest routes). The ASrank method balances between searching the space and sticking to the best solution [8].

#### C. Max-Min Ant System

Ants in the Max-Min Ant System (MMAS) place pheromones only on the path of the best iterative solution, and every tenth iteration, place pheromones only on the path of the best global solution in all iterations. By alternating pheromone deposition between the best iterative and

global solution, searching is avoided only around the best iterative solution, which could be far from the global solution.

MMAS uses the upper and lower limit of pheromones to prevent evaporation of any solution. After updating the pheromones, the MMAS adjusts the values of the pheromones on the grids of all found solutions, whereby if the pheromone value on the grid falls below the lower limit, the pheromone value is adjusted to the lower limit and if the pheromone value on the grid exceeds the upper limit, the pheromone value is adjusted to the upper limit value. Whenever a new best global solution is found, the upper and lower limits of pheromones are updated.

If MMAS begins to stagnate, the values of pheromones on the grids of all solutions will be smoothed, while the best solutions will not lose their relevance and worse solutions will be promoted by increasing the amount of pheromone [9].

#### D. Ant Colony System

Modification of the Ant Colony System (ACS) differs from the basic AS method in the selection of the grid for ant movement as well as in the storage of pheromones. ACS method has local update of pheromones in space. Each ant adjusts the value of the pheromone locally on each grid it passes through during the search.

Using a local pheromone update while searching for a solution improves the search, as each subsequent ant in an iteration knows its predecessors paths, while in a global pheromone update, ants know about other ants paths in the next iteration (after the search of all ants is complete).

After the search of all ants is complete, a global update of the pheromones will be performed. In the ACS method, the pheromone stores only the best ant (the ant that has found the global shortest path in all iterations) [6].

### IV. OTHER OPTIMIZATION APPROACHES.

To compare the pros and cons of PSO and ACO two other optimization approaches were used.

#### A. Diffusion Model

Improving the search capability of Ant Colony Optimization methods is possible by implementing a diffusion model. The diffusion model stores a smaller amount of pheromone on the grids adjacent to the grids of the found path during the global update of pheromones in space. Thanks to the diffusion model, there is a higher probability of selecting adjacent grids of found solutions and thus there is also a higher probability of finding possible better solutions in the vicinity of already found solutions. The diffusion model stores a smaller amount of pheromones on the grids in the radius  $r$  [10]:

$$D_{ij} = \begin{cases} \gamma * \tau_{ij}(t) * \left(\frac{r - d_{ij}}{r}\right) & \text{if } d_{ij} \leq r \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

where  $D_{ij}$  represents the diffusion on the grid  $ij$ ,  $\gamma$  represents the diffusion parameter,  $\tau_{ij}(t)$  represents the value of pheromones on the grid  $ij$  and  $d_{ij}$  represents the distance between the grid of the found solution and the adjacent grid.

#### B. Potential Fields

Ant Colony Optimization methods search space randomly, unlike Particle Swarm Optimization methods, which makes ACO methods not always able to find solutions in the initial iterations. Artificial potential fields compensate this problem and accelerate convergence. The artificial potential fields are calculated by summing the attractive force and the repulsive force of each grid in space [11]:

$$U_{total}(X) = U_{attractive}(X) + U_{repulsive}(X) \quad (7)$$

The attractive force attracts the ant to the target and the repulsive force repels the ant from obstacles. The attractive force decreases by the distance from the target grid and is defined by the following equation:

$$U_{attractive}(X) = \frac{1}{2} * k * p^2(X, X_g), \quad (8)$$

where  $k$  represents the scaling value and  $p$  represents the absolute value of the distance between the grid and the target grid. The repulsive force is distributed at a distance  $p_0$  from the obstacles:

$$U_{repulsive}(X) = \begin{cases} \frac{1}{2} * \eta * \left( \frac{1}{p(X, X_0)} - \frac{1}{p_0} \right)^2 & \text{if } p(X, X_0) \leq p_0, \\ 0 & \text{otherwise} \end{cases}, \quad (9)$$

where  $\eta$  represents the scaling value and  $p$  represents the distance of the grid from the obstacle.

## V. EXPERIMENTS

Experiments of PSO and ACO methods were performed on an environment map of 28x38 grids with densely placed obstacles, which are represented by purple grids. The yellow grids represent the free passage space, the blue grid represents the starting position, and the red grid represents the target (searched) grid, as can be seen in Fig. 1.

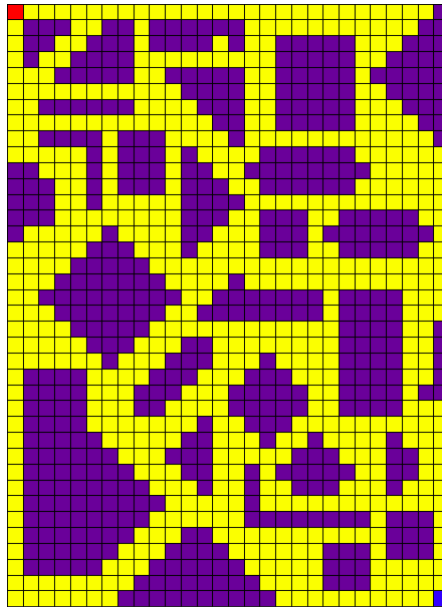


Fig. 1. An example of experimental environment map.

Table explanatory notes:

- Average path lengths – APL
- Average iteration of the first path found – AIFPF
- Average iteration of finding the best path – AIFBP
- Average time execution in seconds – ATE (s)
- Average number of founded paths in seconds – ANFP

Table 1  
PSO experiment

Method	APL	AIFPF	AIFBP	ATE (s)
PSO 0-4	56.2843	<b>1</b>	1	9.1283
PSO 1-3	56.2843	<b>1</b>	1	8.9966
PSO 2-2	56.2843	<b>1</b>	1	8.9811
PSO 3-1	55.8843	<b>1</b>	7.6	8.8702
PSO 4-0	55.2132	<b>1</b>	5	13.6446
TVAC-IW	55.8843	<b>1</b>	38.6	9.9690
SPSO 0-4	55.4843	<b>1</b>	7	12.0691
SPSO 1-3	55.0843	<b>1</b>	34.8	11.9404
SPSO 2-2	54.8500	<b>1</b>	14.2	12.3682
SPSO 3-1	<b>54.5671</b>	<b>1</b>	54.4	11.7612
SPSO 4-0	54.6843	<b>1</b>	19.4	11.9821
QDPSO	56.1269	<b>1</b>	64	19.5638
BPSO	56.2843	<b>1</b>	1	<b>7.9087</b>

The SPSO method found the shortest paths, specifically the SPSO method with a stronger cognitive coefficient. The longest paths were found by the methods QDPSO, BPSO and basic

PSO with a focus on the social coefficient and basic PSO with balanced values of acceleration coefficients. The least computationally intensive method was the BPSO method and the most computationally demanding method was QDPSO.

Table 2  
 ACO experiment

Method	APL	ANFP	AIFPF	AIFBP	ATE (s)
AS	54.9873	1824.2	<b>1</b>	33.4	171.9883
EAS	54.0298	1655	<b>1</b>	<b>19</b>	170.1950
ASrank	54.1269	1884.2	<b>1</b>	19.8	169.5402
MMAS	<b>51.9328</b>	1902.4	<b>1</b>	63.6	88.2875
ACS	53.5612	<b>1969</b>	<b>1</b>	36.2	<b>59.6155</b>

The MMAS method found the shortest paths on average and the basic AS method found the longest paths. The ACS method found the most paths on average and the EAS method found the fewest paths, with the fastest convergence. The MMAS method had the slowest convergence. Approximately as computationally demanding were the AS, EAS and ASrank methods. The ACS method had the least computational complexity.

## VI. CONCLUSION

The results of the experiments show that all PSO and ACO methods were able to find solutions. If it is more important to find the path in the shortest possible time than to find the shortest path, it is more appropriate to use PSO methods. For better results, it is also advisable to use a smaller number of particles as in the case of experiments where twenty particles were used. In this variant, the Standard PSO 2011 method excelled in particular with a higher value of the cognitive coefficient compared to the social coefficient. In case it is more important to find the shortest possible path regardless of time, it is more appropriate to use ACO methods with potential fields, diffusion model and with a higher value of the heuristic factor compared to the pheromone factor. In this variant, the ACS and MMAS methods excelled.

## REFERENCES

- [1] Cleghorn, Christopher W., and Andries P. Engelbrecht. "Particle Swarm Variants: Standardized Convergence Analysis." *Swarm Intelligence*, vol. 9, no. 2-3, 2015, pp. 177–203., doi:10.1007/s11721-015-0109-7.
- [2] Arumugam, M. Senthil, et al. "A Novel Effective Particle Swarm Optimization like Algorithm via Extrapolation Technique." *2007 International Conference on Intelligent and Advanced Systems*, 2007, pp. 515-521., doi:10.1109/icias.2007.4658442.
- [3] Zambrano-Bigiarini, Mauricio, et al. "Standard Particle Swarm Optimisation 2011 at CEC-2013: A Baseline for Future PSO Improvements." *2013 IEEE Congress on Evolutionary Computation*, 2013, pp. 2337-2344., doi:10.1109/cec.2013.6557848.
- [4] Sun, Jun, et al. "Particle Swarm Optimization with Particles Having Quantum Behavior." *Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No.04TH8753)*, 2004, pp. 325-331., doi:10.1109/cec.2004.1330875.
- [5] Yao, Jingzheng, and Duanfeng Han. "Improved Barebones Particle Swarm Optimization with Neighborhood Search and Its Application on Ship Design." *Mathematical Problems in Engineering*, vol. 2013, 2013, pp. 1–12., doi:10.1155/2013/175848.
- [6] Dorigo, M., and L.m. Gambardella. "Ant Colony System: a Cooperative Learning Approach to the Traveling Salesman Problem." *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, 1997, pp. 53–66., doi:10.1109/4235.585892.
- [7] Negulescu, Sorin C., et al. "Elitist Ant System for Route Allocation Problem." *8th WSEAS International Conference on APPLIED INFORMATICS AND COMMUNICATIONS (AIC'08)*, 2008, pp. 62–67.
- [8] Tamura, Yoshiki, et al. "Ant Colony Optimization Using Common Social Information and Self-Memory." *Complexity*, vol. 2021, 2021, pp. 1–7., doi:10.1155/2021/6610670.
- [9] Stützle, T., and H. Hoos. "Improvements on the Ant-System: Introducing the MAX-MIN Ant System." *Artificial Neural Nets and Genetic Algorithms*, 1998, pp. 245–249., doi:10.1007/978-3-7091-6492-1\_54.
- [10] Ji, Junzhong, et al. "A New Mechanism of Pheromone Increment and Diffusion for Solving Travelling Salesman Problems with Ant Colony Algorithm." *2008 Fourth International Conference on Natural Computation*, 2008, pp. 558-563., doi:10.1109/icnc.2008.453.
- [11] Matoui, Fethi, et al. "Local Minimum Solution for the Potential Field Method in Multiple Robot Motion Planning Task." *2015 16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, 2015, pp. 452-457., doi:10.1109/sta.2015.7505223.

# Výkonová analýza prístupových technológií 5G sietí v simulátore NS-3

<sup>1</sup>Maroš BAUMGARTNER, <sup>2</sup>Jozef JUHÁR, <sup>3</sup>Ján PAPAJ

<sup>1,2,3</sup>Katedra elektroniky a multimediálnych telekomunikácií,  
Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach,  
Slovenská republika

<sup>1</sup>maros.baumgartner@tuke.sk, <sup>2</sup>jozef.juhar@tuke.sk, <sup>3</sup>jan.papaj@tuke.sk

**Abstrakt** — V dnešnej dobe sú čoraz väčšie požiadavky na efektívnosť prenosu dát. Zvýšenie efektivity prenosu dát môže byť zabezpečené pomocou 5G sietí a služieb, ktoré ponúkajú. V tomto článku simulujeme služby LTE, enhanced Mobile Broad Band a enhanced Mobile Broad Band with millimeter waves v scenároch s rôznym počtom koncových užívateľov a analyzujeme dosiahnuté výsledky oneskorenia, priepustnosti a jitter. Dosiahnuté výsledky boli získané pomocou sieťového simulátora NS-3. Tento článok dôkladne popisuje a implementuje modul mmWave pre NS-3, pretože je nevyhnutnou súčasťou tejto i budúcej práce. Príspevok končí ilustráciou získaných výsledkov a predstavuje možné témy pre budúcu prácu.

**Kľúčové slová** — mmWave, eMBB, 5G, NS-3, simulácia

## I. ÚVOD

Siete novej generácie známe ako 5G siete sa pomaly stavajú súčasťou nášho života. V budúcnosti prinesú nové príležitosti a rôzne pohľady na dnešnú sieť. Tieto siete sa musia vyrovnáť s rôznymi problémami a výzvami, ktoré vznikajú kvôli stále väčšiemu rozvoju komunikačných charakteristík. V súčasnej dobe je LTE najrozšírenejšia služba na prenos dát v sieti. S postupným narastaním počtu užívateľov v sieti v podobe mobilných telefónov, notebookov a iných prenosných zariadení sa zvyšujú aj nároky na samotnú sieť. Zdravotná starostlivosť, bezpečnosť, logistika, automobilový priemysel a prenos dát, sú oblasti v ktorých musia 5G siete spĺňať prísne požiadavky od koncových užívateľov.

Tieto požiadavky sú napríklad oneskorenie, priepustnosť a kolísanie oneskorenia. V oblasti zdravotnej starostlivosti a bezpečnosti prenosu dát existujú prísne požiadavky na latenciu a oneskorenie siete. Rýchlosť prenosu je nevyhnutná na podporu širokej škály dátových prenosov až do niekoľkých Gbps. Na podporu veľkého počtu koncových používateľov je potrebná škálovateľnosť a flexibilita.

Na základe týchto požiadaviek sme sa v tomto článku zamerali na porovnanie kolísania oneskorenia a priepustnosti starších LTE sietí s technológiami 5G sietí v podobe enhanced Mobile BroadBand (eMBB) a eMBB s milimetrovými vlnami (mmWave).

V prvej časti článku detailne popisujeme blokovú schému a princíp fungovania mm-wave modulu pre simulátor ns3, ktorý sme implementovali a použili pri simuláciách a analyzovaní výsledkov. V tejto časti sú tiež popísané technológie mmWave a eMBB, ktorým je venovaná veľká pozornosť, pretože sú základom tejto a budúcej práce.

V druhej časti popisujeme simulačné scenáre, použité parametre a analyzujeme dosiahnuté výsledky simulácií. Porovnali sme priemerné hodnoty oneskorenia, kolísania oneskorenia a priepustnosti služby LTE s výsledkami získanými simuláciou 5G sietí s implementovaným modulom.

Milimetrové vlny a eMBB budú v budúcnosti základnými službami, ktoré sa budú využívať na prenos dát v 5G sieťach, hlavne kvôli nízkym hodnotám oneskorenia, veľkým prenosovým rýchlostiam a šírke frekvenčného pásma. Šírka frekvenčného pásma je v súčasnej dobe jeden z najväčších problémov, ktorý môžu milimetrové vlny vyriešiť.

## II. TECHNOLÓGIA MILIMETROVÝCH VLN

Nasledujúca časť popisuje technológiu mmWave, eMBB a modul pre sieťový simulátor NS-3. Základ tvorí skôr vytvorený modul pre NS-3 LENA, ktorý je považovaný za najviac robustný [1][2][3].

### A. Milimetrové vlny

Milimetrová vlna je druh elektromagnetickej technológie používanej v rôznych zariadeniach, ako sú telefóny, automobily, lekárske prístroje a ďalšie, na zabezpečenie bezdrôtovej širokopásmovej komunikácie pri vyššej rýchlosti. Toto spektrum obklopené mikrovlnnými a infračervenými vlnami sa dá použiť na vysokorýchlostnú bezdrôtovú komunikáciu, tak ako to umožňuje najnovší štandard Wi-Fi 802.11ad (pracujúci na 60 GHz). Je to extrémne vysoká frekvencia (EHF), pretože predstavuje spektrum signálu od 30 GHz do 300 GHz s vlnovou dĺžkou od 1 do 10 mm. Technológia milimetrových vln navyše poskytuje bezpečnosť komunikačného prenosu [12].

### B. mmWave modul

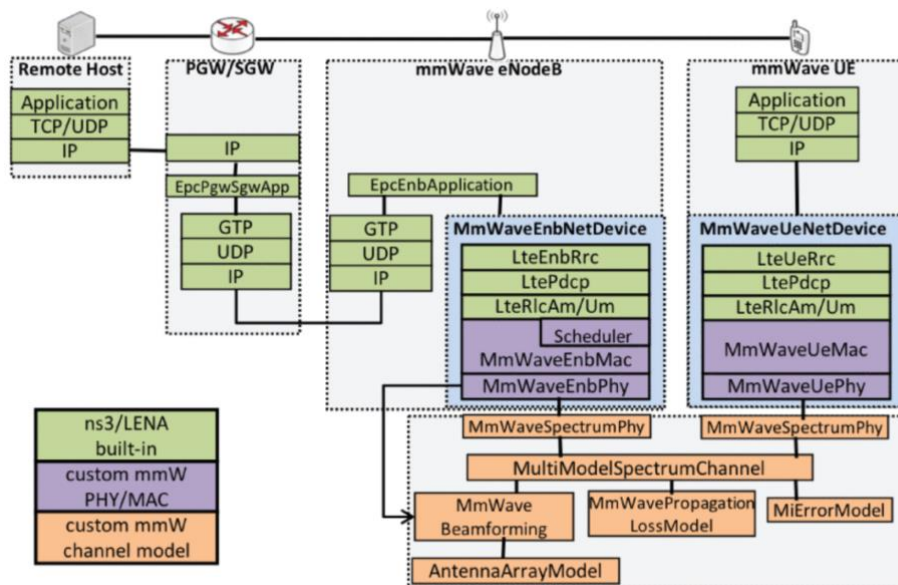
Modul mmWave je navrhnutý na simulovanie komplexných mobilných sietí štýlom 3GPP (3rd Generation Partnership Project) [1][8]. Implementuje LTE / EPC protokoly a vlastné vrstvy PHY a MAC [2][3]. Tieto vrstvy podrobne popísali Rebato a Mezzavilla vo svojej práci [10][11].

Na Obr. 1 je zobrazený diagram tried kompletného mmWave modulu, ktorý obsahuje triedy MmWaveEnbNetDevice a MmWaveUeNetDevice. Tieto triedy slúžia ako rádiové zásobníky mmWave eNodeB a mmWave UE [14]. Celý triedový diagram reprezentuje end-to-end štruktúru simulátora.

Súčasťou tohto modelu je aj trieda McUeNetDevice. NetDevice je zariadenie, ktoré sa vie pomocou duálneho zásobníka pre mmWave a LTE pripojiť k obom technológiám.

MmWaveNBmAC a MmWaveUeMac sú triedy vrstvy MAC. Tieto triedy implementujú užívateľské rozhranie, SAP (Service Access Point) a LTE modul kvôli spolupráci s LTE RLC vrstvou [14].

Do tried MmWaveMacScheduler, odvodených tried a MAC tried je implementovaná podpora pre triedy UM (Unacknowledge Mode), TM (Transparent Mode), AM (Acknowledged Mode) a SM (Saturation Mode), ktoré patria do skupiny RLC.



Obr. 1 Triedový diagram mmWave modulu v NS-3

Kvôli konfiguráciám na vrstve LTE RRC (Radio Resource Control) sa do triedy LteEnbRrc implementuje SAP pomocou MAC scheduler [6][14]. Kvôli tomu je pre vytvorenie spojenia EPC (Evolve Packet Core) požadovaný každý komponent [8][9].

Hlavnou úlohou tried zo skupiny MmWavePhy je mať na starosti riadiace kanály, ktoré sú založené na kontrolných MAC správach a smerový prenos a príjem dát cez downlink a uplink. Každá inštancia triedy Phy komunikuje cez SpectrumChannel pomocou inštancie triedy MmWaveSpectrumPhy [14]. Kvôli fyzickej vrstve modulu mmWave, ktorá je založená na TDD

(Time Division Duplexing) sa trieda MmWaveSpectrumPhy a jej inštancia zdieľa pre uplink a downlink [1]-[6].

Inštancie triedy MmWaveSpectrumPhy do seba zapúzdrujú všetky modely fyzickej vrstvy:

- MmWaveInterference vypočítava interferenciu
- MmWaveSinrChunkProcessor vypočítava signál k interferencii a šum (SINR)
- MmWaveMiErrorModel vypočítava chybový model, ktorý je založený na spoločných informáciách (MI)
- MmWaveHarqPhy vypočítava entitu fyzickej vrstvy a pravdepodobnosť chyby paketu HARQ (Hybrid Automatic Repeat reQuest) [1]-[6][14].

### C. Enhanced Mobile BroadBand

EMBB je definovaný ako rozšírenie existujúcich 4G širokopásmových služieb. EMBB pravdepodobne prináša priepustnosť podobnú vláknu s viac Gbps, zameriava sa na 5G a je štandardizovaná v 3GPP Rel-15 a novších [5]. Môže byť nasadený na frekvenciách milimetrových vln, napríklad v rozsahu 24 - 30 GHz alebo v rozmedzí do 6 GHz, čo je priaznivejšie pre spojenia mimo priamej viditeľnosti, ktoré sa bežne vyskytujú v mobilných komunikáciách. Najširšie pásmo sub-6 GHz definované 3GPP pre 5G NR (nové rádio) je n79, siahajúce cez 900 MHz od 3,3 do 4,2 GHz [6]. Použitie šírky pásma RF 200 MHz, využitie priestorovej rozmanitosti 4 × 4 MIMO (viacnásobný vstup, viacnásobný výstup) a modulácia údajov pomocou 256-QAM (kvadráturňa amplitúdová modulácia) umožňuje dosiahnuť dátovú rýchlosť zhruba 4,6 Gbps - tj. priepustnosť podobná vláknu. Tieto vlastnosti pozitívne ovplyvňujú a zlepšujú robustnosť dát, rýchlosť prenosu, oneskorenie, priepustnosť a samotný prenos dát [7].

## III. SIMULAČNÉ SCENÁRE A PARAMETRE

Táto časť obsahuje popis simulačných parametrov a simulačných scenárov nastavených pomocou simulátora NS-3.

Pre účely výskumu v oblasti robustného prenosu dát pomocou 5g sietí sme sa v tejto práci zamerali na simulovanie a porovnanie jednotlivých metód prenosu dát. V práci porovnáваме a analyzujeme výsledky simulácií LTE, eMBB a mmWave prenosových metód, ktoré boli vytvorené pomocou ns3. Výsledky analyzujeme z pohľadu oneskorenia, kolísania oneskorenia a priepustnosti siete pri rôznych typoch scenárov. Tento typ novej služby 5G sme si vybrali hlavne preto, lebo má prísny a dobre definovaný štandard.

### A. Simulačné parametre

V tejto štúdií bol zvolený prípad, ktorý sa týkal rôzneho počtu koncových užívateľov, ktorý primali dátové pakety. Simulované scenáre prebiehali v simulačnej oblasti o veľkosti 600 m x 600 m, ktorá predstavovala reálne prostredie bez prekážok. Ako transportný protokol bol použitý protokol UDP (User Datagram Protocol) s veľkosťou paketu 512 bajtov. Každý simulačný scenár s rôznym počtom koncových užívateľov, ktorý sa náhodne pohybovali v simulovanej oblasti bol spustený 100 krát.

Tabuľka 1 ukazuje parametre a hodnoty použité počas výskumu.

Tabuľka 1 Simulačné parametre

Parameter	Hodnota
Frekvencia	2,160 GHz (LTE), 6 GHz (eMBB), 30 GHz (mmWave)
Šírka pásma	20 MHz
Počet koncových užívateľov	10, 20, 30, 40, 50
Mobilita koncových užívateľov	Náhodná
Rýchlosť koncových užívateľov	0.277-1.389 mps
Rozmery simulovanej oblasti	600 m x 600 m
Simulačný čas	100 s
Počet simulácií	100
Transportný protokol	UDP
Veľkosť paketu	512 bajtov

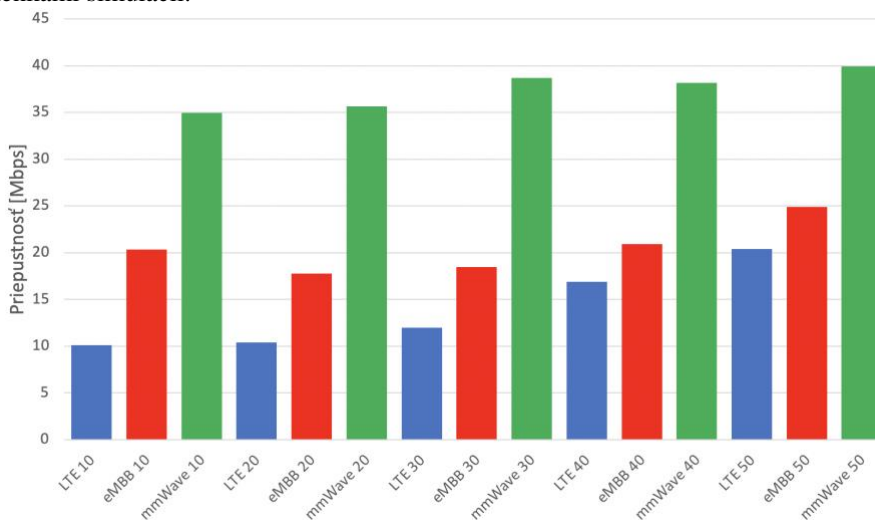
### B. Dosiahnuté výsledky

Konkrétne táto časť obsahuje výsledky výskumu založené na oneskorení, kolísaní oneskorenia a priepustnosti siete. Každý výsledok bol porovnaný s LTE, eMBB a eMBB s mmWave. Bola teda vykonaná štúdiá o kapacite RAN (Radio Access Network) pre tri uvažované rádiové



technológie, ktorých priemerné výsledky sú znázornené v grafoch a tabuľkách nižšie.

Na Obr. 2 sú znázornené priemerné hodnoty priepustnosti pre všetky druhy simulovaných scenárov. Priepustnosť simulovaných služieb sa zvyšuje s narastajúcim počtom uzlov v sieti, ale stojí za zmienku, že z hľadiska služieb každého koncového užívateľa služby eMBB a mmWave výrazne prekonávajú staršiu LTE vo všetkých nami uvažovaných metrikách výkonu s rovnakými podmienkami simulácií.



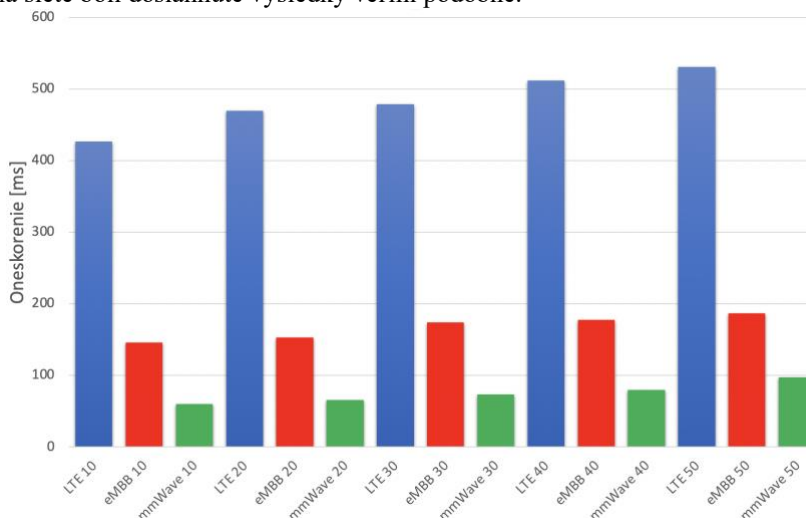
Obr. 2 Priemerné hodnoty priepustnosti pre všetky scenáre

Môžeme vidieť, že priemerné hodnoty priepustnosti pre LTE službu sa pohybujú od 9.7 do 18.4 Mb/s v závislosti od počtu koncových užívateľov a simulovaného scenára, to je výrazne menej ako pri službách 5g sietí. Služba eMBB dosahuje najvyššie úrovne priepustnosti pri 50 uzloch, 22.64 Mb/s. Služba mmWave dosahuje hodnoty v priemere od 33.31 do 39.99 Mb/s čo je výrazne viac ako pri staršej LTE alebo eMBB službe. Tieto dosiahnuté výsledky sú znázornené v Tabuľka 2.

Tabuľka 2 Namerané hodnoty priepustnosti pre všetky scenáre

	10 Koncových užívateľov	20 Koncových užívateľov	30 Koncových užívateľov	40 Koncových užívateľov	50 Koncových užívateľov
<b>LTE</b>	9,7	10,08	12,64	16,88	18,4
<b>eMBB</b>	19,54	18,44	19,24	20,82	22,64
<b>mmWave</b>	33,31	35,76	37,49	38,32	39,99

Na Obr. 3 sú znázornené namerané hodnoty oneskorenia pre všetky typy scenárov. Z pohľadu oneskorenia siete boli dosiahnuté výsledky veľmi podobné.



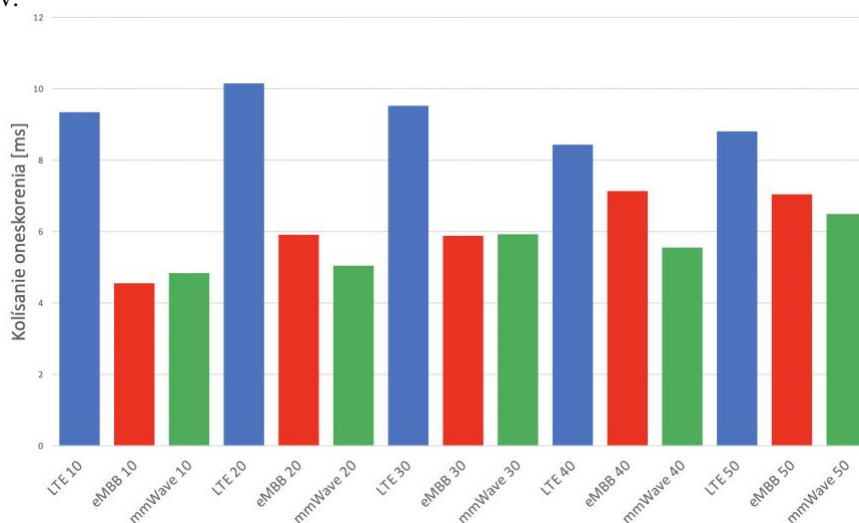
Obr. 3 Priemerné hodnoty oneskorenia pre všetky scenáre

Pri staršej LTE službe boli namerané hodnoty oneskorenia rádovo vyššie ako pri novších eMBB a eMBB s mmWave službách. Priemerné hodnoty oneskorenia služby LTE sa pohybovali v rozmedzí od 426.4 do 530.87 ms pri plne obsadenej sieti. Priemerné hodnoty latencie služby eMBB sa pohybujú v rozmedzí 146.15 – 186.38 ms a eMBB s mmWave pod 100 ms. Tieto výsledky sú znázornené v *Tabuľka 3*.

Tabuľka 3 Namerané hodnoty oneskorenia pre všetky scenáre

	10 Koncových užívateľov	20 Koncových užívateľov	30 Koncových užívateľov	40 Koncových užívateľov	50 Koncových užívateľov
<b>LTE</b>	426,4	469,81	478,87	511,98	530,87
<b>eMBB</b>	146,15	153,04	173,68	177,65	186,38
<b>mmWave</b>	60,2	65,95	73,67	79,85	97,64

Lavacca vo svojej práci [13] dosiahol podobné výsledky pri porovnaní staršej služby LTE s eMBB. Na *Obr. 4* sú znázornené namerané hodnoty kolísania oneskorenia pre všetky typy scenárov.



Obr. 4 Priemerné hodnoty kolísania oneskorenia pre všetky scenáre

Aj pri meraní kolísania oneskorenia služba LTE dosahuje najvyššie úrovne. Priemerné hodnoty tejto meranej veličiny dosahujú hodnoty v závislosti simulovaného scenáru od 8.81 do 10.15 ms. Menšie rozdiely v nameraných hodnotách sú v službách eMBB a eMBB s mmWave kde sa hodnoty líšia priemerne o 1 ms pre každý simulovaných scenár. Tieto výsledky v ms sú znázornené v *Tabuľka 4*.

Tabuľka 4 Namerané hodnoty kolísania oneskorenia pre všetky scenáre

	10 Koncových užívateľov	20 Koncových užívateľov	30 Koncových užívateľov	40 Koncových užívateľov	50 Koncových užívateľov
<b>LTE</b>	8,81	8,44	9,53	9,34	10,15
<b>eMBB</b>	4,56	5,92	5,88	7,14	7,05
<b>mmWave</b>	4,84	5,05	5,93	5,56	6,5

#### IV. ZÁVER

V tejto práci sa zaoberáme komunikáciou cez milimetrové vlny v 5g sieťach s dôrazom na porovnanie tejto služby so službou enhanced Mobile Broadband a staršou LTE v oblasti priepustnosti a oneskorenia siete.

Túto štúdiu sústreďujeme na mmWave preto, že v blízkej budúcnosti bude mať mmWave najväčší vplyv na prenosové siete a rádiový prístup z hľadiska šírky pásma a rýchlosti prenosu pretože mmWave má oveľa širšie prenosové pásmo ako eMBB alebo staršia služba LTE.

Hlavnou analýzou tohto príspevku je vyhodnotenie priepustnosti a latencie novej technológie 5g v porovnaní so sieťou 4g. Jednotlivé scenáre sme simulovali pomocou simulátora ns3, obohatili sme ho modulom na simulovanie 5g sietí pomocou milimetrových vln. Na základe toho

sme boli schopný simulovať jednotlivé scenáre a analyzovať dosiahnuté výsledky. Na základe dosiahnutých výsledkov a výsledkov, ktoré boli analyzované v predchádzajúcej práci sme zistili, že priemerné hodnoty priepustnosti a oneskorenia sú pri službách 5g siete oveľa lepšie ako pri staršej službe LTE. Priemerné hodnoty priepustnosti, oneskorenia a kolísania oneskorenia dosahujú lepšie hodnoty pre eMBB a eMBB s mmWave v porovnaní s LTE. To znamená, že 5G siete sa ukazujú ako efektívnejšie z hľadiska nameraných hodnôt a efektívnosti dátového prenosu ako služba LTE.

Tieto zistenia a simulácie nám pomôžu a budú tvoriť základ budúcej práce v ktorej sa zameriame na obohatenie schopností nášho simulátora začlenením algoritmov a metód, so zvláštnym dôrazom na robustnosť siete a machine learning.

#### POĎAKOVANIE

Výskum prezentovaný v tomto príspevku bol podporený Ministerstvom školstva, vedy, výskumu a športu Slovenskej republiky v rámci výskumného projektu VEGA 1/0753/20, KEGA 009TUKE-4/2019, a Slovenskou agentúrou pre výskum a vývoj s projektom APVV-17- 0208.

#### LITERATÚRA

- [1] N. Baldo, M. Miozzo, M. Requena-Esteso and J. Nin-Guerrero, "An open source product-oriented LTE network simulator based on ns-3", Proc. 14th ACM Int. Conf. Model. Anal. Simulat. Wireless Mobile Syst., pp. 293-298, 2011, [online] Available: <http://doi.acm.org/10.1145/2068897.2068948>, in press.
- [2] TE-EPC Network Simulator, Feb. 2012, [online] Available: [http://iptechwiki.cttc.es/LTE-EPC\\_Network\\_Simulator\\_\(LENA\)](http://iptechwiki.cttc.es/LTE-EPC_Network_Simulator_(LENA)), unpublished.
- [3] H. Tazaki et al., "Direct code execution: Revisiting library OS architecture for reproducible network experiments", Proc. 9th ACM Conf. Emerg. Netw. Exp. Technol. (CoNEXT), pp. 217-228, 2013, [online] Available: <http://doi.acm.org/10.1145/2535372.2535374>, in press.
- [4] The LENA ns-3 LTE Module Documentation, Jan. 2014, [online] Available: [http://iptechwiki.cttc.es/LTE-EPC\\_Network\\_Simulator\\_\(LENA\)](http://iptechwiki.cttc.es/LTE-EPC_Network_Simulator_(LENA)), unpublished.
- [5] M. M. A. Rojas, I. Demirkol, "Analysis of TCP performance for LTE-5G Millimeter Wave Dual Connectivity", UPC, Barcelona, Oct. 2019, [online] Available: <https://upcommons.upc.edu/bitstream/handle/2117/174568/Merly+Asuncion+Rojas+Thesis.pdf?jsessionid=E00D80A41D0CCBE0CECF13D340ECFAF4?sequence=4&fbclid=IwAR2bUWfvx8QraKc81OBOMGfodIf6xGldESnqnVu4vlaVFTmp6Xdkxcp4GEU>.
- [6] A. A. Oliveira, D. Batista, R. Hirata Jr, "Exploring the ns-3 mmWave module", April 2019, [online] Available: [http://vision.ime.usp.br/~arturao/mmwave/mmWaveReport.pdf?fbclid=IwAR3ubQX9V7hrcC546eOvavTNXXBuU9TsMdPjnLy7BKRbXNb54s\\_pizqMkLxI](http://vision.ime.usp.br/~arturao/mmwave/mmWaveReport.pdf?fbclid=IwAR3ubQX9V7hrcC546eOvavTNXXBuU9TsMdPjnLy7BKRbXNb54s_pizqMkLxI).
- [7] S. Dutta, M. Zhang, M. Mezzavilla, M. R. Akdeniz, S. Rangan, "Millimeter wave module for ns-3 network simulator", Polytechnic School of Engineering, New York, [online] Available: [https://www.nsnam.org/wp-content/uploads/2015/04/WNS3\\_2015\\_submission\\_41.pdf?fbclid=IwAR2zhK\\_CLWhg47w6S\\_pXkJrwYzv8qrfVThrvjmxXk99icNaxzm7aXW ntRkI](https://www.nsnam.org/wp-content/uploads/2015/04/WNS3_2015_submission_41.pdf?fbclid=IwAR2zhK_CLWhg47w6S_pXkJrwYzv8qrfVThrvjmxXk99icNaxzm7aXW ntRkI).
- [8] S. Pratschner, et al., "Versatile mobile communications simulation: the Vienna 5G Link Level Simulator", EURASIP Journal on Wireless Communications and Networking, 2018 [online] Available: <https://doi.org/10.1186/s13638-018-1239-6>, in press.
- [9] G. Minopoulos, G. Kokkonis, K. E. Psannis, Y. Ishibashi, "A Survey on Haptic Data Over 5G Networks", International Journal of Future Generation Communication and Networking, pp. 37-54, 2019 [online] Available: <http://dx.doi.org/10.33832/ijfgcn.2019.12.2.04>, in press.
- [10] M. Rebato, M. Zorzi, "Simulation analysis of algorithms for interference management in 5G cellular networks using spatial spectrum sharing", Università Degli Studi di Padova, pp. 17-33, 2015.
- [11] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, S. Rangan, "5G mmWave Module for ns-3 Network Simulator", NYU Polytechnic School of Engineering, New York, June 2015, in press.
- [12] Cesar I., Sandeep R., "The fundamentals of millimeter wave sensors", Texas Instruments, July 2020 [online] Available: [https://www.ti.com/lit/wp/spyy005a/spyy005a.pdf?ts=1616803949996&ref\\_url=https%253A%252F%252Fwww.google.com%252F](https://www.ti.com/lit/wp/spyy005a/spyy005a.pdf?ts=1616803949996&ref_url=https%253A%252F%252Fwww.google.com%252F).
- [13] F. G. Lavacca, P. Salvo, L. Ferranti, A. Speranza, L. Costantini, "Performance Evaluation of 5G Access Technologies and SDN Transport Network on an NS3 Simulator", May 2020 [online] Available: <https://doi.org/10.3390/computers9020043>, in press.
- [14] Edoardo B., "Design and performance evaluation of mm-wave vehicular networks", Politecnico di Torino, pp. 38-44, March 2019.

# Controlling plant growing condition using machine learning

*Marek BERKA, Ondrej KAINZ, Miroslav MICHALKO, František JAKAB*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

marek.berka@student.tuke.sk, ondrej.kainz@tuke.sk, miroslav.michalko@tuke.sk, frantisek.jakab@tuke.sk

***Abstract*** – In our research, we control the environment through machine learning using random forest regression and monitor the ambient temperature and soil moisture. The solution itself is a greenhouse with sensors used to monitor the properties of the environment and components for its modification. The greenhouse supports automatic and manual control mode. The automatic mode of adjusting the temperature and humidity of the greenhouse environment is implemented using the very machine learning. The software solution includes a web application that serves as a user interface for displaying current and previous states of the greenhouse and for setting target values for environmental modifications. The later research is to include incorporate unmanned aerial vehicle both, for security purposes and detection of the growth progress.

***Keywords*** – automated monitoring, environmental influencing, artificial intelligence, machine learning

## I. INTRODUCTION

Crop production has a significant place in the national and supranational economy. Due to the ever-increasing population, the demand for food products to cover the basic necessities of individuals is also growing. Lack of nutritious food is a current problem in many parts of the world. The basis of agricultural production is the environment and soil with suitable conditions for crop growth. Arable land is a limited raw material and therefore there is an effort to make the most of the smallest possible area. We can achieve this through biological, chemical and physical routes, such as the use of genetically modified crops, the most efficient use of fertilizers and pesticides, soil enrichment and, in our case, the optimization of plant living conditions with the help of information technology. Its goal is to increase the profitability of production and optimize growing conditions, which affects the quality, fertility of the plant and the growth rate, and this will help us achieve multiple turnover during one growing season.

Advanced information and communication technologies are entering various areas of industry and are beginning to have a place in agriculture. They enable easy remote monitoring and adjustment of the environment. Another advantage is the possibility of exchanging data within one farm between several seasons or between individual farms and businesses. Until now, information technologies aimed at collecting and evaluating data in agriculture have not been widespread, and therefore data describing ideal conditions for growing individual types of plants and crops are not known. There is also a lack of data on the control of individual components needed to modify the environment. When we reach a certain volume of data, ideally as large as possible, we get space for the potential use of machine learning. In this way, we can achieve process improvements, further data collection and dissemination, which in turn leads to further improvements and better optimization. In our paper we will monitor ambient temperature and soil moisture and try to control it through ventilation and irrigation using machine learning.

## II. MACHINE LEARNING USES IN AGRICULTURE

There are several ways to improve and streamline the process of growing plants with the help of artificial intelligence. One of the problems facing gardeners and farmers around the world is the lack of information. For example, they do not know the right time to plant a particular crop or do not know the right conditions that a crop requires to maximize its quantity and quality.

They lack a nutritional system for recognizing crop deficits well in advance that would allow them to adjust them appropriately. This problem can be solved by collecting and sharing data on growing conditions. By collecting and sharing such data, it is possible to train various solutions for artificial intelligence and neural networks. These will, over time, be able to streamline the cultivation process, such as deciding on the ideal amount of water and light to which the plant should be exposed to maximize yields, which could be predicted[1] using certain machine learning techniques.

However, not only information about the ideal growing conditions of the plant, but also about its current state may not be sufficient. Plants can be endangered by various undesirable insects, pests or diseases that are transmitted by air, water or soil. By using cameras and sensors to monitor the plant and its environment, it is possible to detect threats in real time with the help of artificial intelligence and image recognition. Such monitoring will significantly reduce the time needed for farmers to inspect and make the inspection process itself more efficient, as human error will be avoided and more detail can be captured compared to the human eye. This is also supported by the result of a study[2] where image analysis, using artificial intelligence, was able to distinguish a carrot crop from weeds with almost 75% success without human intervention.

#### *A. Irrigation*

Using various artificial intelligence techniques, it is possible to improve the irrigation process and the application of pesticides and herbicides[3], for example by remote monitoring of soil moisture, automated irrigation by GPS, spraying of pesticides and herbicides by drones while monitoring the condition of plants. The use of autonomous robots increases weed removal efficiency and reduces the amount of crops destroyed in the process.

#### *B. Quality control*

Obtaining cultivation data will increase the quality of cultivated crops. MIT researchers used machine learning to determine growth conditions that would maximize the concentration of volatiles responsible for the taste of basil[4]. Basil was grown under hydroponic conditions in modified shipping containers in Middleton, Massachusetts. Temperature, light, humidity and other environmental factors inside the containers could be regulated automatically. The researchers tested the taste of the plants by looking for certain compounds by gas chromatography and mass spectrometry. The resulting data were inserted into machine learning algorithms developed at MIT and at Cognizant. Research has shown that exposing plants to light 24 hours a day causes the highest concentration of these substances, and thus the best taste of herbs. The research group plans to study how this technology could improve plants ability to fight disease, as well as how different flora can respond to the effects of climate change.

#### *C. Plant classification*

They may be used in the classification systems for orange, papaya, almond, potato, lemon, wheat, maize, rice and soya [5]. The use of these systems provides simple and objective analysis of samples and produces accurate descriptive data. Through these systems it is possible to automate strenuous tasks, in a non-destructive way, to obtain adequate data for future analysis.

#### *D. Disease detection*

Diseases affecting the deciduous parts of crops result in enormous damage to crop production. By using machine and deep learning, it is possible to automatically detect these diseases [6], which makes it possible to prevent the spread of diseases early, which reduces the damage caused.

### III. DESIGN

The aim of this work is to control the process of irrigation ventilation through machine learning in order to maintain constant values of ambient temperature and soil moisture to optimize plant growth. The ventilation and irrigation process takes place by adjusting the environment with the help of fans and a water pump. By using machine learning, it is possible to control these components automatically and achieve automatic maintenance of the targeted value of the environmental properties. Control of components will consist in predicting their performance and work time. Data from the monitored environment collected during temperature adjustment and irrigation would be used to train the regression machine learning model. Since labeled data will be used to train the model, we will use supervised machine learning. The data collected for training model to control ventilation includes:

- temperature before the fans start,
- temperature after the fans finish,
- power of the fans,
- running time of the fans,
- hour of day.

For controlling irrigation, collected training data includes:

- soil moisture before irrigation,
- soil moisture after irrigation,
- power of the pump,
- running time of the pump,
- average temperature since the last irrigation,
- average humidity since the last irrigation,
- hour of day.

#### IV. IMPLEMENTATION

Two machine learning models for ventilation and irrigation control have been developed. The ventilation control model is used to calculate the required percentage power and running time of the fans. The irrigation control model is used to calculate the percentage of pump power and irrigation time. Since in both cases it is necessary to obtain numerical values, these are regression problems. A random forest regressor was used for the implementation of the models. Models for ventilation and irrigation control were trained using the data described in design III. For ventilation control model, the start, end temperature and hour of the day represent the input variables  $X$  and the percentage power of the fans and the running time of the fans represent the output variables  $y$ . For irrigation control model, soil moisture before and after irrigation, temperature, ambient humidity and hour of the day represent the input variables  $X$  and the percentage pump power and irrigation time represent the output variables  $y$ .

#### V. TESTING

Machine learning models for ventilation and irrigation were tested. Testing was performed in an isolated environment measuring 45 cm × 22 cm × 35 cm for 48 hours. The temperature was adjusted with the help of two fans, with a diameter of 80 mm, placed on the side walls of the environment in a push-pull position. Soil moisture was adjusted using the water pump. The ambient temperature targets were set at 23 °C and the soil moisture target was set at 80%. The created regression models of machine learning controlled the power and running time of fans and pumps. The measured values of the ambient temperature during the whole testing period are shown in figure 1, and soil moisture in figure 2. Maximum, minimum and average values of ambient temperature and soil moisture are shown in table 1. From our measured data it follows that the deviations of the data measured during the ventilation control process were lower and thus closer to the entered target value. The deviations of the measured values of soil moisture during the measurement period were higher, and thus further away from our set target value. The resulting average value of measurements, for the entire testing period, is similar to the target value for both processes.

	Target	Minimum	Maximum	Average
Temperature [°C]	23	22.55	24	23.26
Soil moisture [%]	80	73.27	84.45	80.33

Table 1  
Ventilation and irrigation measured values table

#### VI. ON THE FUTURE RESEARCH

In the future research we plan the greenhouse to be powered by renewable resources, primarily the sunlight (solar panel). This should provide independence and self-sustainability. Moreover, the Unmanned aerial vehicle (UAV) should be expected to be used to check the condition of the greenhouse and also the growth progress of the plants. UAV will be also used to monitor the surroundings, used to enhance the security.

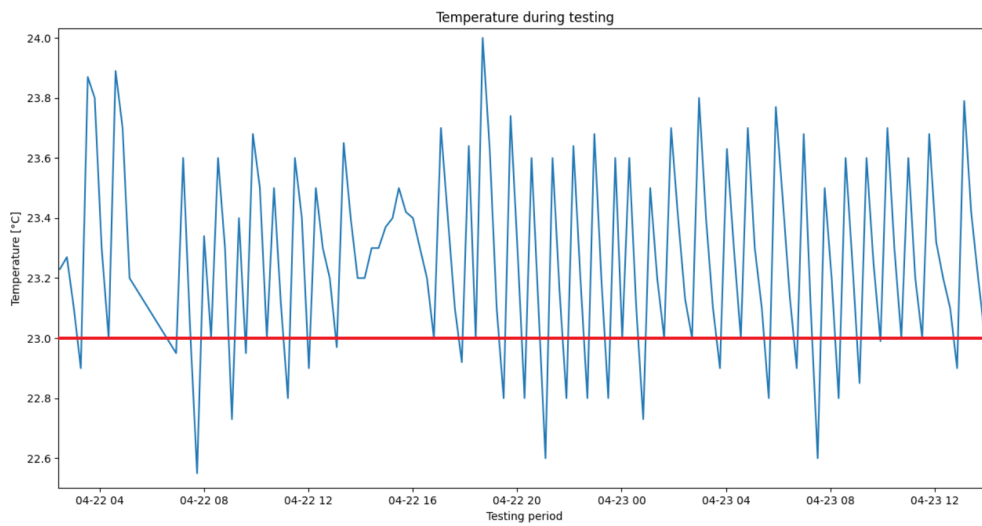


Fig. 1 Measured temperatures during testing period.

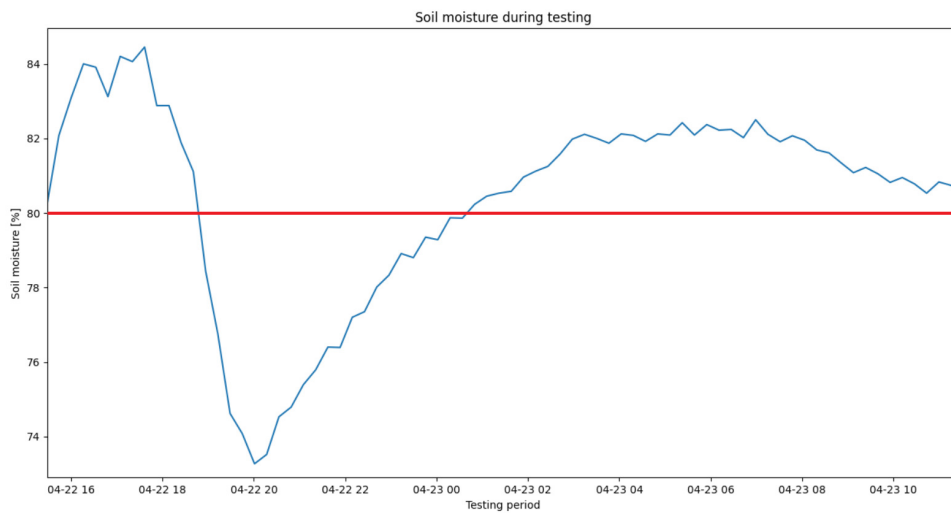


Fig. 2 Measured soil moistures during testing period.

## VII. CONCLUSION

The use of artificial intelligence to improve the cultivation of plants and crops is gradually increasing with the spreading phenomenon. This is due to the constant increase in the number of people, which means the need for faster and more efficient crop cultivation as well as rapid advances in modern technology and increasing simplicity of data collection and sharing.

We were able to use artificial intelligence to control the ventilation and irrigation process in order to maintain ideal environmental conditions for plant growth. The controlled conditions were ambient temperature and soil moisture. The temperature was maintained by fans and soil moisture by a water pump. The power and running time of the devices was controlled using machine learning random forest regression. The success rates in process control varied. Ventilation was able to maintain the ambient temperature with slight deviations from the target value, but during irrigation, deviations of the soil moisture value from the target moisture were more pronounced. The average ambient temperature and soil moisture were close to the target values, and by obtaining additional data, it is possible to increase the accuracy of the created machine learning models.

## ACKNOWLEDGMENT

This publication was realized with support of the Operational Programme Integrated Infrastructure in frame of the project: Intelligent systems for UAV real-time operation and data processing,

code ITMS2014+: 313011V422 and co-financed by the European Regional Development Fund.

## REFERENCES

- [1] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and Electronics in Agriculture*, vol. 151, pp. 61–69, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169917314710>
- [2] M. Aitkenhead, I. Dalgetty, C. Mullins, A. McDonald, and N. Strachan, "Weed and crop discrimination using image analysis and artificial intelligence methods," *Computers and Electronics in Agriculture*, vol. 39, no. 3, pp. 157–171, 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169903000760>
- [3] T. Talaviya, D. Shah, N. Patel, H. Yagnik, and M. Shah, "Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides," *Artificial Intelligence in Agriculture*, vol. 4, pp. 58–73, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S258972172030012X>
- [4] A. J. Johnson, E. Meyerson, J. de la Parra, T. L. Savas, R. Miikkulainen, and C. B. Harper, "Flavor-cyber-agriculture: Optimization of plant metabolites in an open-source control environment through surrogate modeling," *PLOS ONE*, vol. 14, no. 4, pp. 1–16, 04 2019. [Online]. Available: <https://doi.org/10.1371/journal.pone.0213918>
- [5] D. I. Patricio and R. Rieder, "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review," *Computers and Electronics in Agriculture*, vol. 153, pp. 69–81, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169918305829>
- [6] R. Sujatha, J. M. Chatterjee, N. Jhanjhi, and S. N. Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection," *Microprocessors and Microsystems*, vol. 80, p. 103615, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0141933120307626>



# Smart IoT security solution for premises monitoring

*Jozef ELIÁŠ, Ondrej KAINZ, Miroslav MICHALKO, František JAKAB*

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

jozef.elias.3@student.tuke.sk, ondrej.kainz@tuke.sk, miroslav.michalko@tuke.sk, frantisek.jakab@tuke.sk

**Abstract** – The goal of the research described in this paper is to propose smart IoT solution for the monitoring of the premises. The first part of this paper deals with the analysis of the well known communication protocols used in communication between multiple devices. We will cover the three communication IoT protocols MQTT, CoAP and AMQP. Furthermore, algorithms and methods for face detection and recognition are analyzed. The solution should be able to perform the face detection, thus enabling it to recognize the face. The next part of the paper is devoted to the design of security system for area monitoring. The system includes motion detection in the monitored the area using a connected imaging device. In the next research, such solution will be duplicated, allowing to estimate the position of the person in the image. Also the unmanned aerial vehicle (UAV) will be used to capture high-resolution image of the face based on the real-time position of specific object.

**Keywords** – AMQP, CoAP, face detection, face recognition, Internet of Things, MQTT, Viola-Jones

## I. INTRODUCTION

Nowadays it is increasingly necessary to secure the premises in the form of the installation of various security elements and sensors. These security devices that can communicate with each other can be included among the Internet of Things. Communication between these devices is mediated by communication protocols. These protocols are important for forwarding messages between connected devices.

We can increase our security by adding camera that can monitor our premises. To improve security, it is possible to create a system that can detect motion and based on detection, create images and store them in a database. These images can later be viewed by the user and determined the infiltrator by the face detection and recognition algorithms.

## II. IOT COMMUNICATION MESSAGING PROTOCOLS

This section analyzes the most common communication protocols for sending messages between the IoT devices. MQTT and CoAP were developed for applications to collect data from connected sensors, while the another communication protocol AMQP was developed for applications that require reliable business transactions.

### A. MQTT (Message Query Telemetry Transport)

The MQTT communication protocol is may be considered as one of the oldest M2M communication protocols. Protocol was developed in 1999 by A. Stanford-Clark and A. Nipper. Its architecture is based on a client–server model. Due to its easy and wireless implementation, it is often used in embedded systems that have limited device memory and also a limited amount of power battery [1]. It is the most suitable in a large networks with a many of devices or sensors that of send the measured data to the server.

The MQTT protocol works by sending and receiving data. It is based on the publish/subscribe messaging mechanism [2]. The publish mechanism sends the acquired data to the MQTT Broker. The subscribe mechanism then receives the data from the MQTT Broker. It uses the TCP protocol as the transport protocol.

The MQTT protocol offers three levels of Quality of Service [3]. When QoS level 0 is selected, the protocol ensures that the message is sent a maximum of once for each client, but does not guarantee delivery to the destination. QoS 1 ensures that the sender's message is delivered to the recipient at least once. QoS 2 is used when there is no information about a received or lost packet.

### B. CoAP (Constrained Application Protocol)

The CoAP is a lightweight M2M protocol. It is a IoT communication protocol designed for applications and devices used in IoT. Its mostly used in networks with high congestion. Because its communication with devices takes place via the UDP protocol, it is much more faster compared to protocols using TCP communication.

For communication it uses request/response and resource/observe architecture , which is an alternative to publish/subscribe from the MQTT protocol [4]. The principle of sending and receiving messages is similar to the MQTT protocol. Publisher publishes its data to a specific URI. Subsequently, this data will be sent to each subscriber who is subscribed to a specific URI.

Like the MQTT protocol, CoAP supports two types of sending and delivering messages [5]. The first type is considered to be reliable messaging. This means that the sender is informed whether the sent message has reached the recipient. The opposite of such messaging is unreliable messaging. In this case, the sender sends a message for which no confirmation of receipt is needed.

### C. AMQP (Advanced Message Quering Protocol)

The AMQP is a lightweight M2M protocol that was created in 2003. It is considered to be reliable and secure communication protocol. Like the MQTT protocol, for communication it uses a publish/subscribe architecture [4]. It also utilizes TCP protocol with security technologies such as TLS / SSL and SASL to transmit messages.

Like CoAP, it uses only two levels of Quality of Services, reliable and unreliable [4]. With a reliable messaging format, each message is sent at least once. The recipient then acknowledges the received packet with a return acknowledgment packet. An unreliable message format does not guarantee that the message will be delivered to the recipient.

## III. FACE DETECTION USING VIOLA-JONES

Human eye can relatively easily distinguish a human face from other objects, so from a human point of view, the process of face detection may seem simple. From a computer point of view, this process is much more complicated. The face can have different shapes or different skin colors. Therefore, in this work we will focus on two basic principles of face detection based on the human face characteristic, on which the Viola-Jones algorithm is based. The Viola-Jones algorithm is based on the search for a human face based on Haar cascade.

The Viola-Jones algorithm was introduced in 2001 by a pair of two researchers, Paul Viola and Michal Jones [6]. It was created in Matlab software and it is still one of the most used algorithms for face detection today. Its popularity lies mainly in its low false positive results of the face detection and high processing speed.

To correctly detect the human face from the image, the algorithm uses Viola-Jones Haar cascade [7]. These elements, shown in Figure 1, determine how the features of the human face are determined by rectangular selection. Such detection is possible due to specific features, for example human eyes are darker from a person's nose or forehead. Another technique included in this algorithm is the calculation of the integral image. This integral image calculation is always performed at the beginning of the detection of a particular image and is run only once, which speeds up face detection. The integral image has the same dimensions as the original image. However, each pixel represents the sum of all pixels up and to the left of a particular pixel.

## IV. FACE RECOGNITION ALGORITHM

After detecting a face from in image using the Viola-Jones algorithm, the process of face recognition follows. Prior to this process, it is necessary to prepare database of people to be recognized. Using a convolutional neural network a 128-d vector will be created for each face, which defines a human face in terms of vector numbers. The created vector is then stored in the data structure together with the identifier of the person, which in this case identifier represents the name of the person. The encodings of all the people in the database are stored in *.pickle* file.

If the face from the selected image needs to be recognized, the same coding is created for the selected face using a convolutional neural network into a 128-d vector. Subsequently, this vector

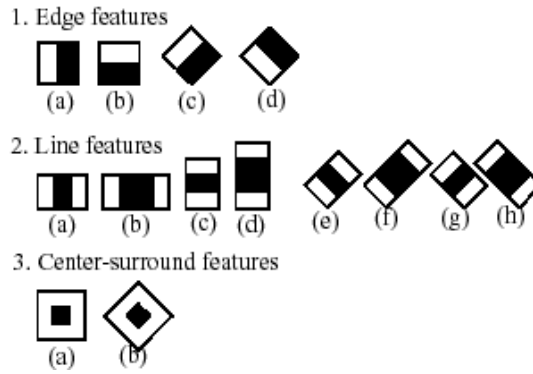


Fig. 1 Haar cascade

is compared with an existing database consists of human faces. If the similarity distance between two vectors are lower than threshold value, system will print the name of the person on user screen. The whole recognition process is shown in Figure 2. In the Figure, the first human face is face from the selected image. The second face is face from the face database. System will calculate 128-d vector form both faces. This vectors are compared. As we can see the distance between this two faces is lower than threshold value so this faces belong to same person.

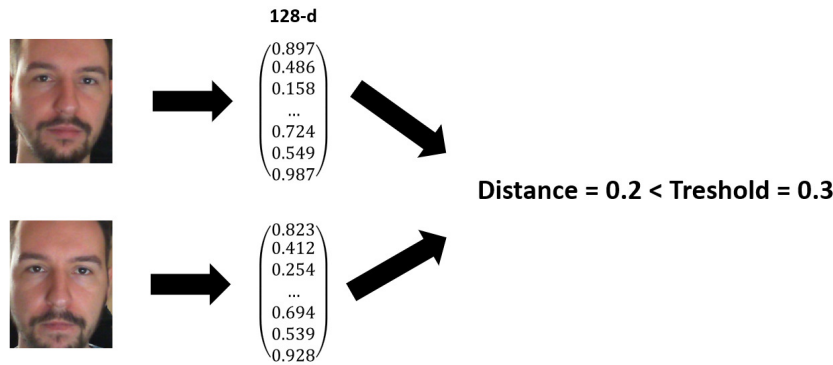


Fig. 2 Face recognition process

## V. MOTION DETECTION ALGORITHM

For system to be called security system must be able to record and detect movement from the area covered by the camera. For motion detection, a hardware motion sensor can be used, which may be placed under the camera. Motion detection is done via comparing of two images. The camera sends the images to the developed software solution. It then compares the current image with the previous image and creates a differential image. The output of this motion detection function is the  $x$ ,  $y$  coordinates and their width and height, which represent the area with the detected motion.

## VI. WEB APPLICATION

For user to communicate with the system, it is necessary to create a web application and a web server on which the application will run. There are several popular options, one of them is the Flask framework. The Flask is a technology based on a template system for the Python programming language. Using the Flask framework of the REST API extension and HTTP methods such as GET and POST, it is possible to get variables from Python into HTML templates and vice versa. Pages are created in Python using paths that are defined in Python language using the `@app.route` ( `^ / ^` ) command. Between the brackets we enter the path of the file that represents the HTML template.

The user connects to the web server using own web browser by typing the IP address of the Raspberry Pi. In the created web solution, the user can watch live video stream from the captured area and also view the created recordings from motion detection.

## VII. SYSTEM DESIGN

The security system for premises monitoring consists of hardware and software parts. The hardware parts of the whole system represent:

- *Microsoft LifeCam HD-3000* - is a web-camera that will be used for capture images from area. The resolution of this camera is 1280x720 pixels.
- *Raspberry Pi 3 Model B* - a minicomputer on which the main functionality of the program will run. It will be used as a web server by using the Flask framework.

## VIII. METHODOLOGY

The task of the system algorithm will be to perform certain steps step by step, at the end of which we obtain the desired goal. This principle of the system is based on performing several steps, including:

- 1) motion detection,
- 2) save the movement record to the database,
- 3) detection of face from the created image,
- 4) calculation of the 128-d vector for the selected face,
- 5) face recognition.

Figure 3. shows the flowchart of the operation of the designed system. When the system is turned on, camera starts sending images to main program for image processing. Based on the motion detection algorithm, it evaluates whether motion has been detected in the scanned area. If motion is detected, the camera takes a snapshot and stores it with the time and date of creation in a database. These records are then selected in a table in the web application. If the user selects a specific record from the table, a face detection algorithm is started. If system finds a face in the image, the face recognition action will start. If the person is in the face database, identity of that person is printed.

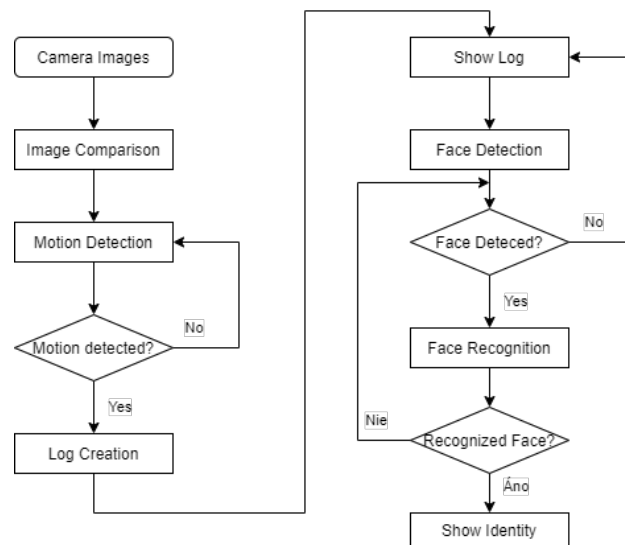


Fig. 3 Flowchart of security system

## IX. ON THE FUTURE RESEARCH

In this research the Viola-Jones algorithm was used. This algorithm is rather old, thus in the next research we plan to test the convolutional neural network with a SSD detector. Also, another imaging device is to be added, thus allowing the estimate the real-time position of the person or the object in the image. Once having position, the unmanned aerial vehicle (UAV) is to capture the higher resolution of the object. The duplicate of hardware solution is expected to be used in the multiple locations, allowing to form smart IoT network for monitoring of larger premises, while also using UAV for high resolution input.

## X. CONCLUSION

The main goal of this research was to design a security solution for monitoring of the premises. The most used communication protocols in IoT were analyzed, including MQTT and CoAP protocols. The system utilizes the Viola-Jones algorithm to detect faces and convolutional neural network to recognize them. When tested in the real conditions, the system was able to correctly detect 15 faces out of 24. However, it also contained a false positives where it detected areas where there was no face. These testing errors were caused by poor lighting of the face and the fact that the faces were not turned directly towards the camera. This drawback of the detection is to be taken into account in the future research, where multiple cameras will be used for the detection using more novel approach along with having as an input high-resolution image of face taken by the UAV.

## ACKNOWLEDGMENT

This publication was realized with support of the Operational Programme Integrated Infrastructure in frame of the project: Intelligent systems for UAV real-time operation and data processing, code ITMS2014+: 313011V422 and co-financed by the European Regional Development Fund.

## REFERENCES

- [1] ERIDANI, Dania; WIDIANTO, Eko Didik. Performance of Sensors Monitoring System using Raspberry Pi through MQTT Protocol. In: *2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. 2018, s. 587–590.
- [2] GUAMÁN, Yesenia; NINAHUALPA, Giovanni; SALAZAR, Gustavo; GUARDA, Teresa. Comparative Performance Analysis between MQTT and CoAP Protocols for IoT with Raspberry PI 3 in IEEE 802.11 Environments. In: *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*. 2020, s. 1–6.
- [3] STANDARD, OASIS. MQTT version 3.1. 1. URL <http://docs.oasis-open.org/mqtt/mqtt/v3>. 2014, roc. 1.
- [4] NAIK, Nitin. Choice of effective messaging protocols for IoT systems: MQTT, CoAP, AMQP and HTTP. In: *2017 IEEE international systems engineering symposium (ISSE)*. 2017, s. 1–7.
- [5] AZZOLA, Francesco. Coapsss protocol: Step-by-step guide. URL: [https://dzone.com/articles/coap-protocol-step-by-step-guide\(26.02.2020\)](https://dzone.com/articles/coap-protocol-step-by-step-guide(26.02.2020)). 2018.
- [6] P. Viola and M. Jones. Robust Real-time Object Detection. *International Journal of Computer Vision*, 57(2):137–154, 2002.
- [7] VIKRAM, K; PADMAVATHI, S. Facial parts detection using Viola Jones algorithm. In: *2017 4th international conference on advanced computing and communication systems (ICACCS)*. 2017, s. 1–4.

# Predikovanie úmrtí na Covid-19 v rôznych časových oknách

*Matúš HUBA, Ján PARALIČ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky  
Technická univerzita v Košiciach  
Letná 9/B, 042 00

matus.huba@tuke.student.sk, jan.paralic@tuke.sk

**Abstrakt** — V tomto príspevku sme sa venovali dátam, ktoré sa týkajú ochorenia Covid-19. Analyzovali sme súčasný stav vo svete a na Slovensku prostredníctvom relevantných vedeckých článkov a zdrojov. Následne sme sa zamerali na denné predikcie úmrtí na Covid-19 na Slovensku v troch rôznych dvojtýždňových testovacích úsekoch, pri troch rôznych dĺžkach tréningových množín. Na modelovanie sme použili modely ARIMA, PROPHET, ETS a TBATS. Výkonnosť všetkých modelov sme merali a porovnávali pomocou strednej kvadratickej chyby (RMSE).

**Kľúčové slová** — Covid-19, modelovanie, analýza, predikcia, časové rady

## I. ÚVOD

Koncom roka 2019 v čínskom meste Wuhan začali hlásiť vysoké počty prípadov so závažným respiračným ochorením. Tento agresívny vírus, nazývaný dočasne ako 2019-nCoV a následne SARS-CoV2, za veľmi krátky čas postihol celé mesto a následne sa rozšíril do iných provincií Číny.

Prvé prípady v Európe začali hlásiť v polovici januára 2020, ale podľa zistení francúzskych epidemiológov sa koronavírus šírila v Európe už v decembri 2019. Po spätnej analýze odobratých vzoriek zistili, že 27. decembra 2019 bol v Paríži hospitalizovaný muž so silným zápalom pľúc, ktorého vzorka bola pozitívna na Covid-19, čo bolo asi 4 dni predtým, kedy Čína oznámila výskyt nového vírusu. Na Slovensku sme prvý prípad SARS-CoV2, čiže pacienta 0, zaznamenali 6. marca 2020. Následne sa v dôsledku nepripravenosti nášho zdravotníctva začali 13. marca 2020 prijímať prísne preventívne opatrenia v zmysle zabrániť šíreniu tohto vírusu.

Prvá vlna najviac postihla najľudnatejšie krajiny, kde kontakt osôb bol priveľký. Už po pár týždňoch od prvého výskytu v týchto krajinách začali hlásiť obrovské množstvo mŕtvych, kedy tento vírus dokázal vyhladiť skoro jednu celú generáciu ľudí v niekoľkých mestách. Naopak menej zaľudnené krajiny a oblasti prežili prvú vlnu bez vysokých čísel úmrtí aj vďaka vládnym nariadeniam a minimalizovaniu kontaktu medzi ľuďmi.

V prípade druhej a niekde aj tretej vlny už to bolo viac-menej vyrovnané vzhľadom na počet obyvateľov jednotlivých krajín a ľudia zomierali aj tam, kde počas prvej vlny do kontaktu s Covidom ani neprišli.

Preto sme sa rozhodli v našom výskume venovať práve analýze dostupných dát o šírení tohto vírusu, ktorý vo svete načasoval novú éru pandémie so zatiaľ neznámym dátumom konca, aby sme dokázali čo najpresnejšie predikovať vývoj najväznejších dôsledkov choroby, t.j. počet úmrtí na najbližšie dva týždne.

Pri našich experimentoch sme využívali dáta z Národného centra zdravotníckych informácií (NCZI)<sup>1</sup>. Naše experimenty sú zamerané na predikciu úmrtí na Covid-19 prostredníctvom štyroch rôznych modelov a to ARIMA, PROPHET, ETS a TBATS. Cieľom tejto štúdie bolo zistiť výkonnosť jednotlivých modelov v rôznych testovacích oknách, čiže rôznych štádiách šírenia Covidu-19 na Slovensku, a tiež analyzovať vplyv rôznej dĺžky tréningového okna použitých časových radov údajov z NCZI.

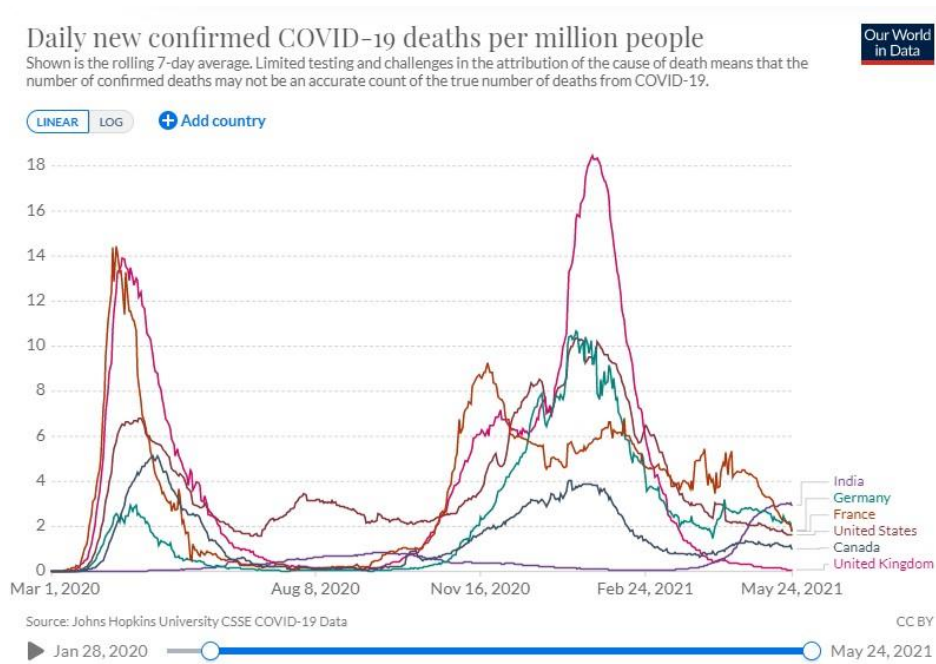
<sup>1</sup> <http://www.nczisk.sk/>

## II. ANALÝZA SÚČASNÉHO STAVU

### A. Analýza denných dát Covidu-19

Tento typ dát je spravidla aktualizovaný každý deň. Cieľom je priniesť občanom aktuálne informácie ohľadom šírenia Covidu – 19 v danom štáte alebo dokonca v konkrétnom regióne, či meste. Informácie v týchto dátach sa zvyčajne týkajú denného prírastku nakazených, počtu testovaných osôb, počtu úmrtí alebo počtu vyliečených ľudí. Takisto v nich nájdeme aj sumarizácie jednotlivých atribútov od začiatku pandémie, alebo počet nakazených na 100 - tisíc obyvateľov. Dostupné sú aj detailnejšie celosvetové dáta<sup>2</sup>, ktoré ponúkajú atribúty ako napríklad vek nakazených a ich pohlavie, počet hospitalizovaných pacientov alebo počet pacientov, ktorí potrebujú umelú pľúcnu ventiláciu.

Tieto dáta zvyknú byť prezentované dvoma spôsobmi. Prvý spôsob je reprezentácia v tabuľkovej podobe a druhým spôsobom je grafická prezentácia, najčastejšie rôzne typy grafov (viď Obr. 1). Zatiaľ čo tabuľky svoju formu nemenia, na druhej strane grafy môžu byť rôzneho typu, napríklad bodové grafy, stĺpcové grafy, koláčové grafy a podobne. Zvyčajne sa jedná o tie isté dáta, ale vyjadrené inou formou.



Obr. 1 Denné prírastky úmrtí na Covid-19 vo vybraných krajinách ku 25.5.2021 (<https://ourworldindata.org/covid-deaths>)

### B. Ostatné analýzy dát Covidu-19

Vedci po celom svete využívajú na zložitejšie analýzy rôzne typy modelov, či už sú to jednoduché regresné, zložitejšie matematické alebo fyzikálne modely. Tieto modely sa využívajú hlavne na zistenie závislostí medzi jednotlivými atribútmi v datasetoch alebo na predikciu vývoja situácie v samotných krajinách a oblastiach.

#### 1) Počítačové simulácie

Na základe lineárnej regresie sa O. Wekesa Simiyu pokúšal zistiť závislosť medzi celkovými potvrdenými prípadmi a celkovými úmrtiami na Covid-19 a následne predpovedať počet úmrtí [1]. Z tohto modelu vyplýva, že čím viac potvrdených prípadov krajina má, tak tým viac má aj úmrtí.

Veľmi zaujímavé modely založené na hlbokom učení použili indickí vedci na krátkodobú a dlhodobú predikciu denného, respektíve týždenného prírastku pozitívnych prípadov v Indii [2]. Na indický dataset aplikovali varianty LSTM modelu ako napríklad konvolučný LSTM, skladaný LSTM a obojsmerný LSTM model. Po vyhodnotení zistili, že obojsmerný LSTM model dokáže krátkodobo predikovať počty pozitívnych prípadov s chybou menej ako 3% a dlhodobo s chybovosťou pod 8%.

<sup>2</sup> <https://ourworldindata.org/>

## 2) *Analýza medicínskych dát*

Keďže toto ochorenie zasahuje pľúca, tak jednou z možností diagnostiky ale najmä sledovania rozsahu poškodenia ochorením Covid-19 je CT pľúc. Jednou z metód bolo vytvoriť 3D modely pľúc nakazených pacientov, aby mohli identifikovať najviac infikované časti pľúc a následne vytvoriť algoritmus na rýchlu diagnostiku Covidu-19 [3].

Jednou z publikovaných analýz je aj analýza tehotných žien pozitívnych na Covid-19 a následne ich novorodencov, kde sa skúmal prenos medzi matkou a plodom [4]. Zo štúdie vyplynulo, že pri 29 novorodencoch ktorí boli testovaní, sa nevyskytli žiadne prípady ochorenia Covid-19, pričom sa následne testovali aj placenty matiek, ktoré boli negatívne na koronavírus. Svetová zdravotnícka organizácia (WHO), ktorá skúmala v Číne 147 tehotných žien, ktoré boli buď pozitívne alebo podozrivé na Covid-19, dospela k záveru, že gravidné ženy nemajú predpoklad na vážny priebeh choroby pri prípadnom nakazení [4].

## 3) *Analýzy dát z diskusií na sociálnych sieťach*

Pandémia koronavírusu rozpútala vo svete veľkú verejnú diskusiu. Názory na tento vírus sú rôzne a preto aj postoje ľudí k tejto situácii sú odlišné. Populárne sú dáta analyzujúce diskusie na sociálnych sieťach [5], konkrétne na sociálnej sieti Twitter, alebo dáta analyzujúce záujem ľudí navštevovať oficiálne stránky ohľadom Covidu-19 pomocou viacnásobného regresného modelu [6].

Pri prvom výskume vedci skúmali mieru sentimentu daných príspevkov a prišli s výsledkom, že skoro polovica príspevkov mala pozitívny sentiment, pričom negatívny sentiment sa vyskytol v 31,1% príspevkoch [5].

Ďalší výskum analyzoval úmysel sledovať oficiálne stránky ohľadom Covidu-19, kde výskumníci na vzorke 120 ľudí zistili, že záujem navštevovať oficiálne stránky je ovplyvnený kvalitou webstránky, prístupom a subjektívnym názorom jednotlivca [6].

## 4) *Analýzy vplyvu a dodržiavania preventívnych opatrení*

Rozdiely v opatreniach medzi európskymi štátmi a ázijskými krajinami, čo sa týka rýchlosti a intenzity opatrení, boli obrovské. Zatiaľ čo viacero ázijských krajín prišlo s okamžitým rozsiahlym testovaním, následným hľadaním kontaktov a ich okamžitou izoláciou, tak v Európe tieto procesy boli značne oneskorené a takisto aj prevenčné opatrenia, ako napríklad nosenie rúšok, sa prijalo v Ázii oveľa lepšie v porovnaní s Európou [7].

Propagačná analýza čínskych vedcov [8] nám podáva obraz o situáciách v niektorých štátoch Ázie a Európy pri prepuknutí Covidu-19, respektíve pri prvých výskytoch infikovaných v týchto štátoch. V tejto analýze porovnávali reprodukčné číslo, dobu inkubácie vírusu a počet denných prírastkov nakazených ako faktory, ktorými sme schopní sledovať šírenie koronavírusu s dátami z čínskej provincie Hubei. Pomocou Gaussovej distribučnej teórie sa snažili predpovedať šírenie ochorenia Covid-19 pred zavedením opatrení a následne vyhodnotili situáciu po ich prijatí. Napríklad 13. marca 2020 pri výskyte 15 000 nakazených v Taliansku, vedci predpokladali do konca marca 250 000 nakazených pri žiadnych opatreniach, ale Taliansko zaviedlo opatrenia 8. marca 2020 a podľa reprodukčného čísla Číny, dosiahne koncom marca počet nakazených číslo 100 000 [8].

Taktiež sme našli dáta analyzujúce vplyv dodržiavania základných opatrení [9] ako dodržiavanie odstupov, nosenie rúšok, poprípade celotvárové masky, či používanie ochranných okuliarov na prenos tohto vírusu. Výsledky tejto štúdie vlastne potvrdili už dnes známe opatrenia za správne a odstupy od ľudí, prekrytie horných dýchacích ciest, či poprípade ochrana očí okuliarmi boli pred vyvinutím vakcín jedinou cestou ako bolo možné s týmto vírusom bojovať na individuálnej úrovni.

Pochopenie takýchto dát môže v budúcnosti pomôcť jednotlivým štátom, inštitúciám alebo jednotlivcom prekonať túto pandémiu ľahšie a oveľa efektívnejšie.

## III. POUŽITÉ METÓDY PREDIKCIE

### A. ARIMA

ARIMA (p, d, q) modely slúžia na predpovedanie časových radov, pričom hodnota parametra „p“ nám hovorí o stupni automatickej regresie, „d“ hodnota hovorí o stupni diferencovania a hodnota „q“ hovorí o stupni kľzavého priemeru pri použití rôznych intervalov v časovom rade. Podmienkou použitia tohto typu modelu je stacionárnosť časového radu. Ak zvolené časové rady nespĺňajú túto podmienku, tak jednou z možností ako to dosiahnuť je diferencovanie cieľových hodnôt. Druhá možnosť je použitie funkcie `auto.arima()`, ktorá cieľové hodnoty diferencuje priamo pri výpočte podľa potreby. Pre modelovanie našich dát sme zvolili funkciu `auto.arima()`, ktorá nám vráti najlepší model, respektíve najlepšie p, d a q hodnoty, podľa koeficientov AIC alebo BIC. AIC aj BIC poskytujú spôsob, ako nájsť v modeli rovnováhu medzi vhodnosťou a zložitnosťou.



### B. PROPHET

PROPHET je metodika predikovania implementovaná v jazykoch R a Python. Je určený na predpovedanie časových radov postavený na aditívnom modeli, kde nelineárne trendy zodpovedajú ročným, týždenným a denným sezónnosťami s „holiday“ efektom. Najlepšie funguje na časových radoch, ktoré majú silné sezónne vplyvy. Taktiež je odolný voči chýbajúcim hodnotám a zmenám v trende.

### C. ETS

ETS metodika slúži na predpovedanie jednorozmerných časových radov a zameriava sa na trendové a sezónne komponenty, ako aj na aditívne chyby.

### D. TBATS

TBATS je metóda vhodná na predpovedanie časových radov, ale so zložitými sezónnymi vzormi prostredníctvom exponenciálneho vyhladzovania.

## IV. PREDIKOVANIE ÚMRTÍ POMOCOU METODIKY CRISP-DM

### A. Pochopenie cieľa

Ako cieľ tejto štúdie sme si určili predikciu úmrtí na Covid-19 na Slovensku v troch 14-dňových testovacích úsekoch od 1.10.2020, 1.12.2020 a 1.2.2021. Na trénovanie našich modelov sme použili 4-týždňovú, 8-týždňovú a 12-týždňovú dĺžku trénovacích dát pred zvolenými dátumami testovania. Rozhodli sme sa predikovať časové rady prostredníctvom štyroch predikčných modelov a to pomocou vyššie predstavených modelov ARIMA, PROPHET, ETS a TBATS. Ich výkonnosť sme merali pomocou strednej kvadratickej chyby (RMSE).

### B. Pochopenie dát

Ako sme už spomenuli, na tieto predikcie sme použili dáta z Národného centra zdravotníckych informácií (NCZI). Konkrétne sa jednalo o dataset denných štatistík, ktorý zaznamenáva denné prírastky pozitívne testovaných pomocou RT-PCR testov a neskôr aj pomocou Ag testov. Nájde tu aj informácie o dennom počte vykonaných RT-PCR a Ag testov, ale obsahuje aj atribút kumulatívneho počtu úmrtí. Číže v ňom evidujeme 7 atribútov, z čoho šesť je numerických a jeden je nominálny (ukážka dát je na Obr. 2). Dataset obsahoval v čase našej analýzy 364 záznamov od 6.3.2020 do 4.3.2021.

	Datum	Pocet.potvrdenych.PCR.testami	Dennych.PCR.testov	Dennych.PCR.prirastkov	Pocet.umrti	AgTests	AgPosit
1	2020-03-06	1	378	1	0	NA	NA
2	2020-03-07	3	66	2	0	NA	NA
3	2020-03-08	5	76	2	0	NA	NA
4	2020-03-09	7	69	2	0	NA	NA
5	2020-03-10	7	111	0	0	NA	NA

Obr. 2 Ukážka z datasetu denných štatistík od NCZI

Pri skúmaní rozloženia hodnôt v jednotlivých numerických atribútoch sme zistili, že hodnoty sú zošikmené napravo, a to z dôvodu pomalého nástupu pandémie na Slovensku, keď sme dlhý čas mali iba minimálne počty nakazených a tomu odpovedala aj minimálna úmrtnosť na tento vírus.

Chýbajúce hodnoty sa vyskytli iba v dvoch atribútoch, a to zhodne po 60,16%. Išlo o atribút, ktorý eviduje denný počet vykonaných Ag testov (AgTests) a atribút, ktorý zaznamenáva denný prírastok pozitívnych pri Ag testovaní (AgPosit). Celkovo sa v dataseete nachádza 17.2% chýbajúcich hodnôt.

### C. Príprava dát

Zo spomenutého datasetu denných štatistík sme si vyseletovali atribút dátumu a atribút kumulatívneho počtu úmrtí na Covid-19 na Slovensku. Keďže sme sa rozhodli predikovať denné prírastky úmrtí, tak sme si vytvorili nový atribút denných úmrtí, ktorý sme dostali diferencovaním hodnôt atribútu kumulatívneho počtu úmrtí.

### D. Modelovanie

Ako prvé sme si vytvorili potrebné časové rady pre trénovacie množiny v už spomenutých trénovacích časových oknách. Taktiež sme si vytvorili testovacie množiny už v spomenutých časových oknách testovania. Dohromady sme natrénovali 36 predikčných modelov časových radov. Po vytvorení týchto modelov sme pre každý jeden model vypočítali jeho výkonnosť, ktorú sme merali pomocou strednej kvadratickej chyby (RMSE).

E. Vyhodnotenie

1) Zhrnutie z hľadiska dĺžky trérovacej množiny

Tabuľka 1 Vyhodnotenie z hľadiska dĺžky trérovacej množiny

Sumár počtu najlepších modelov z hľadiska dĺžky trérovacej množiny			
	1.10.2020	1.12.2020	1.2.2021
4 týždne	3-krát	raz	raz
8 týždňov	0	0	0
12 týždňov	raz	3-krát	3-krát

V Tabuľka 1 môžeme vidieť, že všetky najúspešnejšie modely podľa ich strednej kvadratickej chyby boli natrérované buď na 4-týždňovej trérovacej množine alebo na 12-týždňovej trérovacej množine. Veľkou zaujímavosťou je, že ani jeden z týchto modelov pri jednotlivých testovacích úsekoch nebol najlepší pri 8 týždňoch trérovania.

Najlepšie 4-týždňové modely pri prvej testovacej množine môžeme odôvodniť štruktúrou ich trérovacej množiny, v ktorej sa v porovnaní s ostatnými trérovacími množinami nachádzalo minimálne množstvo nulových hodnôt.

Pri druhom testovacom úseku bolo najviac najúspešnejších modelov trérovaných na 12 týždňovej trérovacej množine, keď predikcie týchto modelov pokračovali v plynulom stúpajúcom trende.

Pri poslednej testovacej množine bolo najviac najlepších modelov taktiež trérovaných na 12 týždňovej trérovacej podmnožine dát, tak ako pri predchádzajúcej testovacej množine. Zistili sme, že TBATS modely boli najlepšie v presne opačných trérovacích podmnožinách ako zvyšné tri modely, či už to bolo pri prvom, druhom alebo treťom testovacom úseku, ktoré sú uvedené v Tabuľka 1 ako najlepšie raz.

2) Zhrnutie z hľadiska modelu

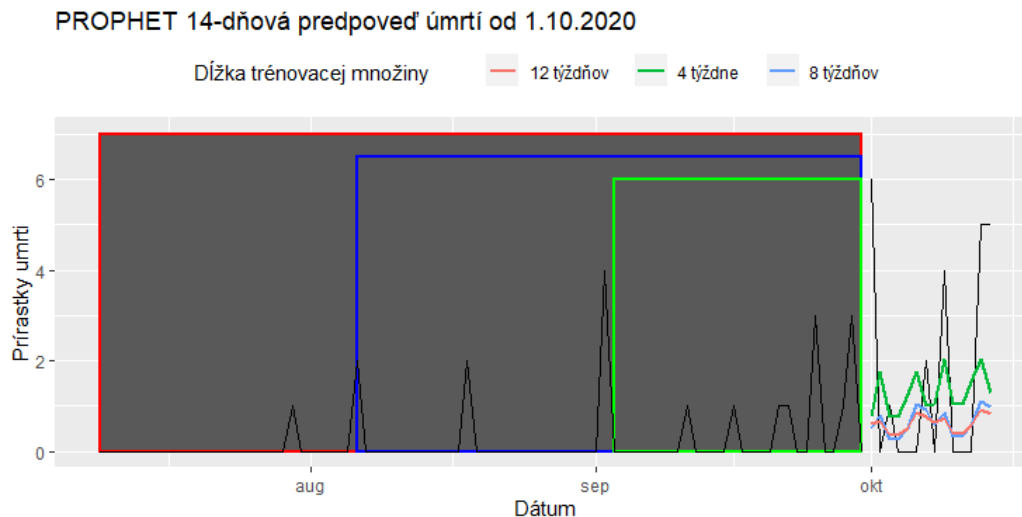
Tabuľka 2 Vyhodnotenie z hľadiska modelu

Zhrnutie z hľadiska modelu			
	1.10.2020	1.12.2020	1.2.2021
4 týždne	PROPHET	TBATS	ARIMA
8 týždňov	PROPHET	ETS	PROPHET
12 týždňov	TBATS	ARIMA	ARIMA

V Tabuľka 2 môžeme vidieť zhrnutie našich najlepších predikčných modelov pri každej trérovacej množine jednotlivých testovacích úsekov podľa ich strednej kvadratickej chyby. Z globálneho pohľadu vidíme, že najfrekventovanejšie modely v tejto tabuľke sú PROPHET a ARIMA modely, každý bol najlepší trikrát. Zatiaľ čo PROPHET bol najfrekventovanejší pri prvom testovacom úseku od 1.10.2020, tak ARIMA zase pri poslednom testovacom úseku od 1.2.2021. Ostatné dva modely sa v tabuľke nachádzajú po dvoch, respektíve po jednom.

Tieto výsledky nám hovoria podľa výsledkov RMSE, že použitie PROPHET modelov je vhodnejšie pri takomto type dát na množinu nižších hodnôt, ako v tomto prípade začiatok druhej vlny pandémie, pričom ARIMA modely sú vhodnejšie na použitie na množinu dát s vyššími hodnotami, čiže vrchol tejto pandémie.

Po vizualizácii jednotlivých modelov (pozri príklad takejto vizualizácie na Obr. 3) sme videli, že niektoré modely pôsobili optimistickým dojmom, keď predikovali v jednotlivých testovacích úsekoch pokles denných úmrtí na Covid-19, iné zase na druhú stranu pôsobili pesimistickým dojmom, keď predikcie stúpali, ale boli aj modely, ktoré predikovali konštantne. Tieto situácie boli samozrejme v korelácii s dĺžkou trérovacej množiny, či jednotlivého testovacieho úseku, kde boli pandemické situácie na Slovensku rozdielne, keď napríklad 1.10.2020 na Slovensku ešte len začínala druhá vlna, pričom 1.12.2020 a 1.2.2021 už bola druhá vlna v plnom prúde a na Slovensku pribúdali desiatky až stovky denných úmrtí na Covid-19.



Obr. 3 Predpoveď počtu úmrtí na Slovensku pri rôznych tréningových dĺžkach od 1.10.2020 pomocou PROPHET modelu

## V. ZÁVER

V tomto článku sme sa venovali analýze dát o ochorení Covid-19, konkrétne predikcii úmrtí na Covid-19 na Slovensku. Na vytvorenie týchto predikcií sme použili dáta z NCZI, pričom úspešnosť jednotlivých predikčných modelov sme merali prostredníctvom strednej kvadratickej chyby (RMSE). Zistili sme, že najfrekventovanejšími najlepšimi modelmi v jednotlivých časových oknách boli ARIMA a PROPHET, pričom všetky najlepšie modely boli tréňované na 4 týždňovej alebo 12 týždňovej tréningovej množine dát.

## POĎAKOVANIE

Táto práca vznikla vďaka podpore Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-16-0213 a Vedeckej grantovej agentúry MŠVVaŠ SR a SAV projekt č. 1/0685/21.

## REFERENCIE

- [1] Wekesa Simiyu, Oliver, Linear regression: Predicting total deaths from total coronavirus cases. April 2020, Preprint, Dostupné na internete: doi: <http://dx.doi.org/10.13140/RG.2.2.36818.91848>
- [2] P. Arora a kol, Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos, Solitons & Fractals, Vol. 139, október 2020, doi: <https://doi.org/10.1016/j.chaos.2020.110017>
- [3] X. Wang a kol, A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization From Chest CT. in IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2615-2625, Aug. 2020, doi: 10.1109/TMI.2020.2995965
- [4] David A. Schwartz, An Analysis of 38 Pregnant Women With COVID-19, Their Newborn Infants, and Maternal-Fetal Transmission of SARS-CoV-2: Maternal Coronavirus Infections and Pregnancy Outcomes. in Arch Pathol Lab Med; vol. 144, no. 7, pp. 799–805, Jul. 2020, doi: <https://doi.org/10.5858/arpa.2020-0901-SA>
- [5] M. Hung a kol, Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence. in J Med Internet Res, vol. 22, no. 8, e22590, Aug. 2020, doi: 10.2196/22590
- [6] S. Sumaedi a kol, A model of intention to use official COVID-19 websites. Dostupné na internete: <https://www.emerald.com/insight/content/doi/10.1108/HE-07-2020-0048/full/html>
- [7] E. Han a kol, Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. Dostupné na internete: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)32007-9/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)32007-9/fulltext)
- [8] L. Li a kol, Propagation analysis and prediction of the COVID-19. Dostupné na internete: <https://www.sciencedirect.com/science/article/pii/S2468042720300087>
- [9] Derek K Chu a kol, Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. The Lancet, Vol. 395, číslo 10242, júl 2020, str. 1973-1987, doi: [https://doi.org/10.1016/S0140-6736\(20\)31142-9](https://doi.org/10.1016/S0140-6736(20)31142-9)

# Vizualizácia publikačnej činnosti výskumného tímu

*Lenka DOBRANSKÁ, Viera PAVLIŠKOVÁ, František BABIČ*

Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Letná 9, 042 01 Košice, Slovenská republika

lenka.dobranska@student.tuke.sk, viera.pavliskova@tuke.sk, frantisek.babic@tuke.sk

**Abstrakt** — Článok popisuje vizualizáciu publikačnej činnosti vybraného výskumného tímu pomocou interaktívnej aplikácie. Jednotlivé vizualizácie boli identifikované na základe komunikácie s členmi daného výskumného tímu. Cieľom je poskytnúť rôzne pohľady na príslušnú dátovú vzorku v jednoducho pochopiteľnej forme, napríklad vývoj počtu publikácií v čase pre daného autora, počet publikácií v jednotlivých kategóriách po rokoch, počet publikácií pre vybraných autorov v čase alebo podľa kategórií, spoluautorské siete (čím je počet spoluprác medzi autormi vyšší, tým je hrana medzi dvoma uzlami, ktoré predstavujú autorov hrubšia). Získané informácie môžu pomôcť autorom pri plánovaní ich budúcej publikačnej činnosti, alebo pri vyhodnocovaní miery spolupráce s rôznymi odborníkmi z rôznych výskumných oblastí.

**KLúčové slová** — Publikačná činnosť, spoluautorská sieť, sieťová analýza, aplikácia

## I. ÚVOD

Hodnotenie vedy z kvantitatívneho ako aj z kvalitatívneho hľadiska je nevyhnutné na zistenie nových trendov, existujúcich medzier vo výskume, nových príležitostí pre spoluprácu medzi autormi, či vplyvu výskumu, ktorý sa môže odrážať na výslednom počte citácií alebo hodnote h-indexu. Príležitosť v tomto prípade predstavuje kontinuálne zhromažďovanie dát o publikačnej činnosti, ktoré následne vytvára dátovú základňu so zaujímavým objemom a potenciálom odhaliť skryté súvislosti a užitočné zistenia.

Komplexné siete sa v posledných rokoch dostávajú do popredia, odkedy účinne popisujú širokú škálu systémov zahŕňajúcu množstvo rôznych disciplín ako biológia, informačné technológie či spoločenské vedy. Za posledné dve desaťročia sa sieťová veda stala novou a veľmi dôležitou disciplínou. Môže sa považovať za nový akademický odbor, nakoľko v roku 2005 americká rada pre výskum definovala sieťovú vedu ako novú oblasť základného výskumu [1].

V teórii grafov vrcholy predstavujú spoločenské entity (v sieťovej vede sa vrcholy nazývajú uzly), ktoré môžu zastupovať širokú škálu jednotlivých subjektov ako ľudia, organizácie, krajiny, výrobky, rastliny či zvieratá; a hrany predstavujú väzby medzi nimi (v sieťovej vede sa priame interakcie medzi nimi nazývajú odkazy) [2]. To znamená, že čiara, ktorá spája dva vrcholy, môže predstavovať množstvo druhov vzťahov ako napríklad komunikácia, spolupráca, priateľstvo, príbuzenstvo, známosti či obchod. Inými slovami, sociálna sieť pozostáva z konečnej množiny vrcholov a vzťahov alebo väzieb [2].

Nadviazané vzťahy môžu mať osobný alebo profesionálny charakter. Okrem sociálnych vzťahov môžu odkazy predstavovať aj tok informácií, tovaru alebo peňazí. Štruktúru takýchto sociálnych sietí zvyčajne znázorňujú grafy, preto sa siete a grafy považujú za rovnocenné.

### A. Analýza Súčasného Stavú

Téme spoluautorstva a korelácií s akademickými výkonmi sa venoval podrobnejšie Qianwen Ariel Xu spolu s Victorom Changom [3]. Študovali vnútornú štruktúru siete spoluautorstva a vzťah medzi sieťou a akademickým výkonom autorov v sieti. Na účely tohto výskumu zhromaždili bibliografické údaje 166 autorov z troch najlepších vysokých škôl v Šanghaji. Analyzovali ukazovatele centrálnosti na určenie najdôležitejších vrcholov v grafoch, aby identifikovali najvplyvnejšie osoby v sociálnej sieti, či egocentrickú efektívnosť siete. Následne použili analýza odkazov a zhlukovú analýzu. Autori vypočítali aj Spearmanov korelačný koeficient s cieľom identifikovať vzťahy medzi akademickým výkonom autorov a metrikami sieťovej analýzy. Autori vďaka analýze dokázali určiť najvýznamnejšieho autora v sieti. Výpočtami zistili požadované vzťahy medzi výkonmi akademikov a metrikami sieťovej analýzy,

čím dokázali identifikovať produktívnych autorov. Výskum naznačuje, že množstvo autorov môže nadviazať užšie vzťahy spolupráce s ďalšími vedcami z iných skupín a taktiež to, že by sa mali vyhnúť spolupráci s väčším počtom autorov z rovnakej skupiny, nakoľko to môže viesť k neefektívnosti.

Nasledujúca štúdia je založená na analýze viac ako 812-tisíc verejných autorských profilov služby Google Scholar [4]. Cieľom autorov bolo vykonať demografickú analýzu a získať najskôr informačný prehľad o autoroch z rôznych aspektov, ako je napríklad krajina, z ktorej pochádzajú, oblasť vedy, ktorej sa venujú, či získané akademické tituly autorov. Na základe publikačných záznamov si ako ďalší cieľ stanovili vytvoriť globálnu spoluautorskú sieť a skúmať koreláciu medzi spoluautorstvom a jednotlivými citačnými metrikami. Po vytvorení globálnej spoluautorskej siete ju autori študovali z rôznych aspektov. Najskôr vykonali demografickú analýzu, neskôr použili metriky sieťovej analýzy, ako je napríklad stupeň, koeficient zhľukovania či PageRank pre lepšie pochopenie siete. Autori sa rozhodli analyzovať aj jednotlivé skupiny vytvorené na základe získaného akademického titulu, či oblasti ich výskumu. V neposlednom rade študovali vzťah medzi sieťovými metrikami a citačnými metrikami pomocou Pearsonovho korelačného koeficientu. Jeden z výsledkov štúdie predstavuje silu korelácií medzi jednotlivými metrikami, napríklad silnú koreláciu medzi ukazovateľom PageRank a H-indexom. Nadobudnuté poznatky dokážu uplatniť v praktických prípadoch, ako je napríklad hodnotenie vedeckého vplyvu jednotlivca alebo inštitúcie.

Ďalšia štúdia skúma spoluautorskú sieť v oblasti scientometrie pomocou metód analýzy sociálnych sietí s cieľom pochopenia výskumnej spolupráce v tejto vedeckej komunite [5]. Autori pracovali s údajmi o spoluautorstve, ktoré získali z internetovej databázy Web of Science 3 024 vedcov, 1 207 inštitúcií, 68 krajín a 22 akademických oblastí. Táto štúdia sa zameriava na štyri vlastnosti siete na makroúrovni, a tými sú hustota, zložka, stredná vzdialenosť či koeficient zhľukovania. Metriky na mikroúrovni sa zameriavajú na hodnotenie vlastností jednotlivých aktérov. Medzi skúmané sieťové metriky na mikroúrovni patrí stupeň, blízkosť či medzil'ahlosť. Na základe vypočítaných sieťových matrik autori zostavili tabuľku s najproduktívnejšími autormi, ktorí majú zároveň najvyšší rozsah spoluprác. Dokázali zistiť, že až 54,78% z celkového počtu článkov sú práve spoluautorské články.

## II. NÁVRH VLASTNÉHO RIEŠENIA

### A. *Vízia Riešenia*

Víziou nášho riešenia bolo vytvorenie interaktívnej webovej aplikácie, ktorá by poskytovala autorom užitočné prehľady a súvislosti medzi jednotlivými záznamami ich publikačnej činnosti, vytvorené na základe analýz dátovej vzorky príslušnej výskumnej skupiny. Aplikácia bude implementovaná vo forme populárneho dashboardu v programovacom jazyku R, ktorý by sme rozdelili do jednotlivých sekcií, nakoľko by cieľovej skupine výskumných pracovníkov pôsobiacich na univerzite poskytovala príjemné a hlavne prehľadné používateľské rozhranie.

Dôležitým krokom pred samotnou implementáciou bola identifikácia funkčných a nefunkčných požiadaviek cieľovej skupiny. Medzi funkčné požiadavky sme zaradili možnosť zvolenia atribútov a konkrétnych hodnôt pri generovaní grafov a spoluautorských sietí, t.j. typ publikácie, rok vydania publikácie, autorské kolektívy, alebo ich kombinácie. Z pohľadu nefunkčných požiadaviek bola dôležitá interaktivita implementovanej aplikácie, možnosť priblížiť alebo oddialiť príslušnú autorskú sieť, rôzne možnosti filtrovania vstupných údajov. Používateľské rozhranie malo byť navrhnuté tak, aby bol s aplikáciou schopný pracovať aj používateľ, ktorý ju vidí po prvý krát.

### B. *Popis Dátovej Množiny a Experimenty*

Spracovali sme vybranú dátovú vzorku, ktorá predstavuje publikačnú činnosť členov Centra hospodárskej informatiky na KKUI FEI TUKE. Dátová vzorka obsahuje celkovo 746 záznamov od roku 1996 do roku 2020, ktoré majú definované napríklad kategóriu, rok vydania, hlavného autora, spoluautorov, atď.

Pomocou prvotných úprav a následných experimentov sme dospeli k nasledujúcim zisteniam:

- Celkový počet unikátnych hlavných autorov predstavuje hodnotu 158.
- Celkový počet unikátnych spoluautorov predstavuje hodnotu 411.
- Celkový počet všetkých unikátnych autorov a spoluautorov predstavuje hodnotu 496.
- Ktorí autori boli v ktorom roku najproduktívnejší.
- Priemerný počet autorov na publikáciu predstavuje hodnotu 3.
- Celkové počty publikácií v jednotlivých kategóriách a v konkrétnych rokoch.
- Najvyššia hodnota 24 opakovaných spoluprác medzi tromi autormi.

- Najčastejšie spolupráce medzi autormi v konkrétnych rokoch a konkrétnych kategóriách.

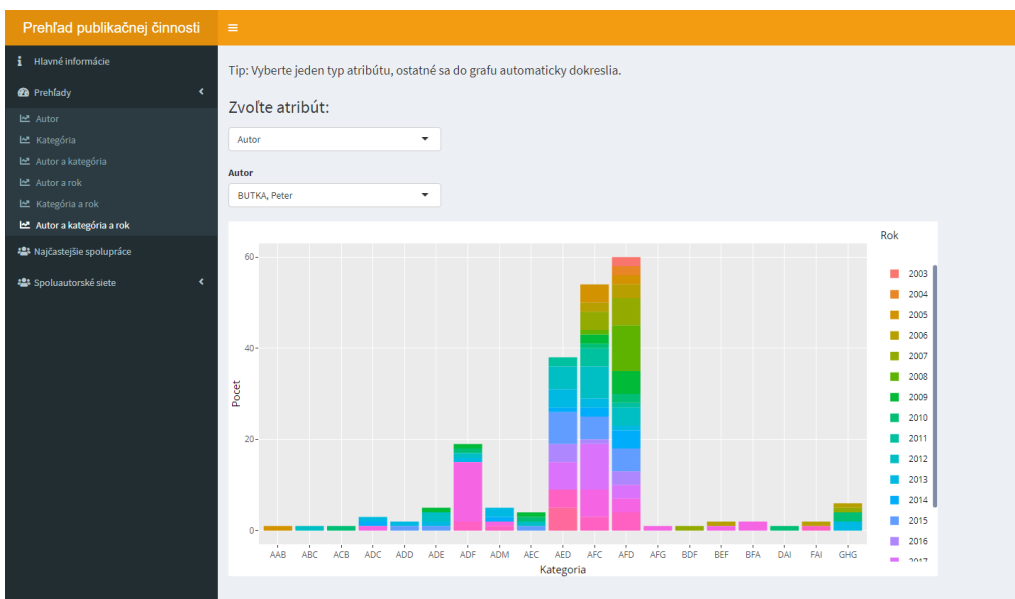
### C. Popis aplikácie

Do prvej sekcie sme umiestnili základné informácie a charakteristiky dátovej vzorky, s ktorou sme pracovali.



Obr. 1 Sekcia so základnými charakteristikami

Do aplikácie sme následne pridali sekciu prehľadov, do ktorej sme umiestnili jednotlivé grafy. Sekciu prehľadov sme ešte rozdelili na podsekcie, pričom do každej sme umiestnili len jeden graf podľa obsahu zobrazených informácií. V každej sekcii sa nachádza jeden alebo viacero filtrov pre zotriedenie informácií, ktoré chce používateľ vizualizovať. Prvá časť je zameraná na jednotlivých autorov, kde si používateľ dokáže zvoliť konkrétneho autora a zobraziť v čase jeho publikačnú činnosť. Druhá sekcia je zameraná na kategórie, kde si používateľ zvolí konkrétnu kategóriu a zobrazí sa mu počet vydaných publikácií v jednotlivých rokoch v danej kategórii. V tretej podsekcii si používateľ zvolí konkrétneho autora a typ publikácie. Nasledujúca podsekcia je venovaná autorovi a konkrétnemu roku. V piatej podsekcii je možné filtrovanie podľa kategórie a roku, pričom v zobrazenom grafe získavame informácie o tom, koľko publikácií vydali autori vo zvolenej kategórii a roku. V poslednej podsekcii si používateľ zvolí najskôr typ atribútu a až tak jeho konkrétnu hodnotu. Zvyšné atribúty sa do grafu dokreslia automaticky. Tento typ grafu poskytuje používateľovi väčší prehľad nad publikačnou činnosťou.



Obr. 2 Podsekcia autor a kategória a rok

Po dokončení prehľadov sme vytvorili sekciu najčastejších spoluprác, kde sme umiestnili tri interaktívne tabuľky. V prvej tabuľke má používateľ možnosť vidieť konkrétne mená autorov a počet ich opakovaných spoluprác. V druhej tabuľke sa nachádzajú najčastejšie spolupráce medzi

autormi v konkrétnych rokoch a v tretej tabuľke počet rovnakých spoluprác v konkrétnych kategóriách.

V poslednej fáze vývoja sme pridali sekciu so spouautorskými sieťami. Vytvorili sme celkovo tri interaktívne spouautorské siete so všetkými autormi, ktoré zobrazujú ich prepojenia. Prvá sieť zobrazuje aj príslušné počty spoluprác (váhy) medzi jednotlivými autormi. Čím je počet spoluprác medzi autormi vyšší, tým je hrana medzi dvoma uzlami, ktoré predstavujú autorov hrubšia. Po úspešnom vygenerovaní spouautorskej siete s váhami sme pridali do aplikácie ešte podsekcie, kde si používateľ môže vygenerovať spouautorskú sieť na základe výberu konkrétneho roku, či kategórie publikácie.



Fig. 3 Sekcia spouautorských sietí

#### D. Testovanie

Príprava na testovanie spočívala zo stanovenia konkrétnych testovacích úloh a vytvorenia dotazníka, v ktorom sa respondenti vyjadria k samotnej práci s aplikáciou a prípadným návrhom zmien. Testovacích úloh sme si pripravili celkovo deväť. Snažili sme sa zostaviť tieto úlohy tak, aby pokryli testovanie každej sekcie v aplikácii. Pripravený dotazník obsahoval celkovo osem tvrdení, ktoré sa týkali spokojnosti s testovaním a s aplikáciou. Tvrdenia sa nachádzali v druhej časti dotazníka a tester mali možnosť vyjadriť svoj súhlas, respektíve nesúhlas s danými tvrdeniami.

Zaujímalo nás napríklad, či sa používateľom zdajú byť vygenerované grafy prehľadné a ľahko pochopiteľné. Celkovo 60% opýtaných súhlasilo s daným tvrdením. Tiež nás zaujímalo, či by používatelia odporučili používanie tejto aplikácie aj iným výskumným skupinám. Až 80% opýtaných s tvrdením súhlasilo.

Tester mali možnosť vyjadriť ďalšie svoje pocity, názory, postrehy, či odporúčania na funkčnosť a používanie aplikácie. Medzi najčastejšie odporúčania k vylepšeniu aplikácie, ktoré sme následne realizovali patrili:

- Zjednotenie škál v jednotlivých vygenerovaných grafoch.
- Pridanie vysvetlení ku skratkám publikácií.
- Zotriedenie zoznamu autorov a skratiek publikácií abecedne.
- Zotriedenie zoznamu rokov vzostupne.
- Pridanie podrobnejšieho popisu k poslednej podsekcii prehľadov.
- Odstránenie legend vo vygenerovaných grafoch, kde legenda a os x zobrazujú totožné informácie.
- Používateľ má po novom možnosť výberu len z konkrétneho zoznamu kategórií, v ktorých vybraný autor publikoval, nie zo zoznamu všetkých kategórií.

### III. ZÁVER

Naším cieľom bolo analyzovať príslušnú dátovú vzorku a vytvoriť interaktívnu aplikáciu, ktorá by vizualizovala publikačnú činnosť konkrétnej výskumnej skupiny. Vďaka vykonanej analýze a následne implementovanej R Shiny aplikácie sme dokázali spätne vizualizovať jednotlivé záznamy o publikačnej činnosti a odhaliť medzi nimi skryté súvislosti. Pomocou interaktívnych grafov, spoluautorských sietí a popisov, dokáže aj bežný používateľ tieto súvislosti jednoducho a rýchlo pochopiť. Prínosom tejto aplikácie sú práve vyplývajúce informácie o kreativite a produktivite autorov, ktoré im v budúcnosti čiastočne môžu pomôcť s plánovaním ich publikačnej činnosti. Vďaka jednotlivým záznamom o interakciách s ostatnými vedcami dokážu zhodnotiť mieru schopnosti získať prístup k novým vedomostiam z rôznych oblastí, čím sa môže zvýšiť ich efektívnosť.

### POĎAKOVANIE

Táto práca vznikla vďaka podpore Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-16-0213 a Vedeckej grantovej agentúry MŠVVaŠ SR a SAV projekt č. 1/0685/21.

### ZOZNAM REFERENCIÍ

- [1] National Research Council, *Network science committee on network science for future army applications*, Washington D.C., USA: National Academies Press, 2005, pp. 14-15.
- [2] S. Wasserman, K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge, UK: Cambridge University Press, 1994, pp. 36-37
- [3] A. X. Qianwen, V. Chang, "Co-authorship network and the correlation with academic performance," in *Internet of Things*, vol. 12, no. 10030, 2020.
- [4] Y. Chen X. Fu, P. Hui, "Building and Analyzing a Global Co-Authorship Network Using Google Scholar Data," *Proceedings of the 26th International Conference on World Wide Web Companion*, pp.1219–1224, 2017.
- [5] M. Erfanmanesh, V. Rohani, A. Abrizah, "Co-authorship network of scientometrics research collaboration," in *Malaysian Journal of Library & Information Science*, vol. 17, no. 3, pp.73-93, 2012.



# Cloud Based System for Autonomous Vehicle Navigation

<sup>1</sup>Ján VAŠČÁK, <sup>2</sup>Dušan HERICH

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

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

<sup>1</sup>jan.vascak@tuke.sk, <sup>2</sup>dusan.herich@student.tuke.sk

**Abstract** – With the emergence of autonomous vehicles a need to steer them effectively and safely arises. This paper deals with designing a remote control system for multiple vehicles with the utilization of cloud technologies using the D\* algorithm. The designed system utilizes ROS2 system and in connection with path planning arranges functionality for localization and map sharing among units. The efficiency of the proposed system was proved by a series of experiments.

**Keywords** – Cloud, D\* path planning, navigation

## I. INTRODUCTION

The present-day emergence of autonomous vehicles renders a necessity for the effective control of such systems, referring primarily to the navigation divisible into three sub-tasks, namely path-planning, mapping, and localization in the environment. Furthermore, resulting from the deployment of multi-vehicle systems, further requirements arose, demanding an efficacious data interchange among vehicles to ease the problem of navigation in a time-variant environment by equipping the system with updated information about its state. Accordingly, this work aims to develop and create a system for navigation permitting the operation of multiple vehicles in dynamic environments, interchanging data utilizing a cloud computing platform.

## II. CLOUD ROBOTICS NAVIGATION SYSTEMS

This section gives an overlook about main functions, which are in navigation and its associated problems. Hence, it addresses mapping, localization, and path planning in cloud robotics environments.

### A. Mapping

The mapping process characterizes data acquisition and deposition performed by the robot at the time of exploration in an advisable representation [1]. In addition, mapping is deemed affianced to the localization emanating from the interdependent nature of the two, meaning the robot is unable to determine its position in an unknown environment, while it can not construct a map without bearing an information of its position. Resulting from the implementation of the mapping algorithms with utilization of cloud computing, several approaches emerged, categorisable as follows:

- Complete offloading,
- Partial offloading,
- Pre-mapping.

The complete offloading of mapping focuses on shifting the computation to the cloud environment completely, therefore relocating the computational burden posed by the algorithms from the robot to the remote computing environment. Experimental results [2] of the described approach indicated throughput of 5MB/s in the network, manifesting the limits of the technique resulted in dropped frames due to the bandwidth limitations.

The alternative approach is a creation of distributed framework [3] for Simultaneous Localization and Mapping (SLAM) partitioning the computation among the remote cloud environment and the robot's computing unit. The system employs two parallel threads for its operation. The first

thread handles map calculation by performing non-linear optimization on a set of key-frames. The other thread evaluates position of the camera on a known map. The described separation rendered the possibility of the map optimization thread to the cloud while sustaining a low network bandwidth requirements.

Particular applications permit omission of map construction by SLAM algorithms [4] by limiting robot's movement to a set of predetermined paths decreasing the computational and network bandwidth requirements representative for the SLAM algorithms to a factual minimum. On the occasion of system's operation in a dynamic environment, the systems employ off-board and on-board perception systems to gather information concerning the pre-mapped environment's present state. Furthermore, the employment of the cloud computing enables gathered information propagation through the fleet of operating vehicles.

### *B. Localization*

The robot localization describes a mobile robot's estimation of position in consideration of the surrounding environment [5]. Adoption of cloud technologies bounded to the concepts of IoT, smart-space and wireless sensor networks facilitated approaches to localization bounded to not only on-board sensor systems such as rotational encoders or accelerometers and gyroscopes.

The feasible approach is utilization of the Radio Frequency IDentification (RFID) technology [6]. The utilization of RFID readers and active tags mounted on the objects necessitating location estimation enabled for implementation of signal attenuation model on the cloud servers, therefore, the object's position is calculated by reader-tag distance determination calculated by the model from the Received Signal Strength Indicator (RSSI) and Link Quality Indicator (LQI) gathered from readers.

Another approach [7] engages in localization estimation by utilization of cloud computing for image recognition. The system operates by processing robot-mounted camera images via the artificial neural network located in the cloud environment, estimating mobile vehicle's position from the known images. The evaluation of the system confirmed its functionality after training the artificial neural network with images of four locations, pointing to the significant disadvantages of the system primarily the requirement for pre-training the neural network and limited position estimation accuracy.

### *C. Path planning*

The path planning is interpreted as the determination of subsequent operations aimed at the vehicle translation from the preliminary to the successive location [8]. Followingly, path planning algorithms ordinarily appeal to the transfer costs minimization, especially in reference to minimal distance trajectory determination.

The fabrication of the Path Planning as a Service model [9] concentrates on facilitation of the on-demand path planning as a cloud service with the minimum-cost path plan calculation serving mobile robots. The system implementing the outlined model comprises three layers:

- client robot layer,
- cloud engine layer,
- server layer.

The primordial layer represents fleet of physical robots and facilitates a web-socket based communication among units and the cloud engine via JSON-encoded messages encompassing information concerning unique vehicle identification and a map location, rendered vital for path planning accomplishments.

The cloud engine layer employs Rapyuta [10] cloud robotics engine devising a computing environment for the respective software execution. The layer is further separated into four components providing load-balancing, action-planning by setting the velocity to the physical robot units, path plan calculation, and map on which the planning is performed.

The server layer constructed upon the Hadoop Distributed Filesystem (HDFS) facilitates data interchange among robot units, primarily path plan informations, map data and unique robot's description.

The system examination embodied the multi-robot people tracking application utilizing two Turtlebot robots employing depth cameras. Trials regarded the robots current position as a starting location, while the depth camera tracking the human torso facilitated the goal location determination. The systems discussed application demonstrates its functionality following a human employing a Dijkstra algorithm utilizing designed Path Planning as a Service system.

### III. SYSTEM DESIGN AND IMPLEMENTATION

This paper's elementary ambition is the construction of the cloud based navigation system for multiple vehicles operating in a dynamic environment utilizing the D\* path planning algorithm, whose conventional schema can be depicted in Fig. 1 [11].

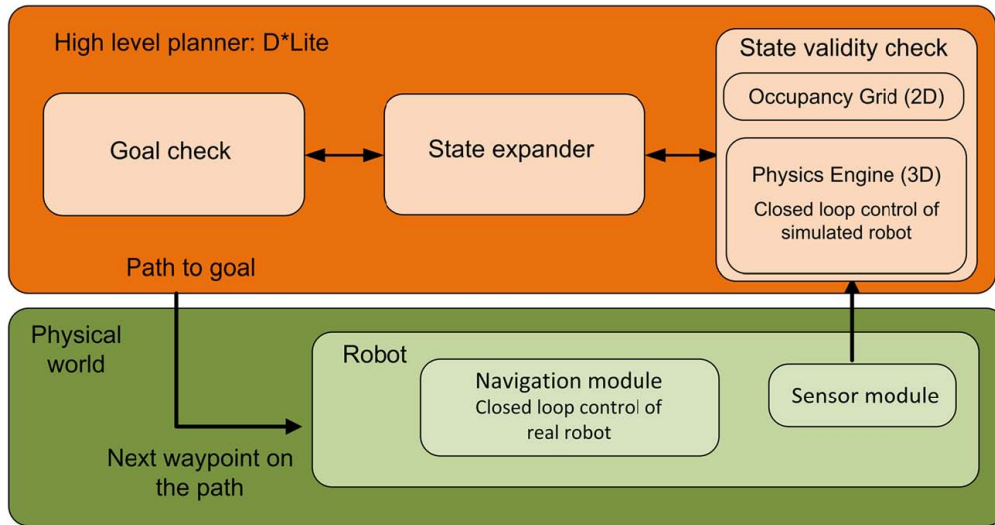


Fig. 1 Block chart of a conventional path planner using D\* algorithm.

However our contribution is based on the development of an architecture using cloud technologies. It is comprised of five layers as demonstrated in Fig. 2 with operation separated to the local and the remote computing environment represented by the Amazon Web Services (AWS) cloud computing platform.

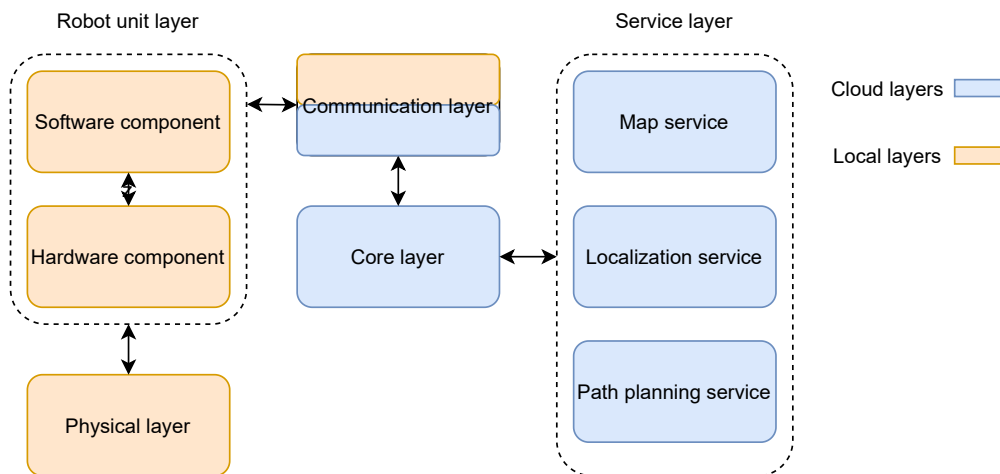


Fig. 2 Designed cloud robotics system architecture.

#### A. Physical Layer

The presented layer abides the lowest level in the designed architecture intended to subsidize the robot layers operation located above, by introducing the structurality to the environment by employing apparatus bolstering the individual units sensing capabilities present in the robot layer by employment of the passive RFID tags distributed regularly on the environment's surface forming a grid with a fixed cell resolution creating a pre-mapped environment.

#### B. Robot Unit Layer

The robot unit layer is a representation of physical mobile robots deployed in the system, separable into two components, particularly a hardware component constituting the electrical equipment installed in the unit and a software component facilitating functionality of the hardware

component. The hardware component comprises the RFID reader accountable for scanning tags placed in the physical layer the vehicle hoovers over, ultrasonic distance sensor for hindrance detection and odometric sensors on every wheel to facilitate position calculation. Furthermore, the hardware component encompasses two electric motors and the Raspberry Pi Zero W computer as a computational unit.

The software component warrants the functionality of the hardware component and is liable for the interaction with the communication layer. Resulting from its purpose, the structure of the software component reproduces the structure of the hardware component, hence separates its modules among sensors and actuators sets. Modules in respective groups are accountable for respective hardware parts. The software layer interacts directly with the connection layer by transmitting immediate sensory readings, respectively receiving settings for actuators from the upper layers in the architecture.

### *C. Communication Layer*

The introduced layer's primary objective is information interchange among robot units and the cloud environment. The layer hinges on the AWS Internet of Thing service managing the vehicle's connection to the cloud environment and accommodating the software permitting the integration to the AWS platform.

For the proper operation, the discussed layer requires an internet connection facilitated by individual computing units in the robot unit layer and a wireless access point, consequently demanding each vehicle to remain within its operational range. Provided compliance with the outlined criteria, the layer initiates communication via the means of Message Queuing Telemetry Transport (MQTT) protocol. Transmitted messages are JavaScript Object Notation (JSON) encoded files, comprehending three types of values, particularly:

- Desired: values calculated by the upper layers signifying the values actuators present in the robot unit should achieve, e.g speed of the particular motor.
- Reported: portray values obtained from the vehicle's sensors and values of current actuator's settings.
- Delta: depicts the difference between desired and reported values.

Additionally, every message contains metadata containing a timestamp and a message number enabling identification of specific messages vital for proper operation of all layers.

### *D. Core layer*

The core layer manages the formation of robot's unit clone in the cloud environment to provision setting for performing computational tasks off-board the robot's processing unit. Structurally, the bestowed layer positions above the communication layer, interacting undeviatingly with it to control the actions performed on the corresponding unit and below the service layer, utilized to carry out a computation if there is a need to weigh in data generated by the system as a whole.

The layer's structure correlates to the robot unit layer's software component enabling core layer's modules assignment to the respective sensor or actuator on the robot unit. Utilization of the Elastic Compute Cloud (EC2) service delineates the virtual server instance employment for creation of clones in one-to-one manner. The layer's operation is based on the Robot Operating System 2 (ROS2) providing algorithms for sensor values evaluation and storage of the current vehicle's state. Additionally, the layer engages in communication with the service layer if there is a necessity to consider data generated by the system as a whole.

### *E. Service layer*

Typifying the uppermost layer, the service layer administers services to the core layer. The module deployed on the layer is obliged to endure two indispensable characteristics to be suitable for deployment in the service layer; specifically, to access data shared among robots operational in the system and a not time-critical assignment for the robots task accomplishment; thus, those obligations prelude modules such as drive controllers from the presence in the layer due to the absence of both characteristics. The layer primarily builds upon the AWS EC2 service to execute deployed services and the AWS Relational Database Service (RDS), representing a managed relational database service in the cloud accommodating data sharing from various units. Additionally, due to the layer utilization by a variant number of robot instances, it needs to comply with cloud computing resemblances, especially elasticity and availability. Therefore, the layer implements load balancers redirecting incoming request among instances considering their utilization and automatic scaling to fluctuate number of available resources by the actual demand generated from the lower layers of the system. Conjunction of the two allows for instance state

monitoring, hence on the occurrence of failure or unavailability, the system automatically ceases operation of the affected instance and replaces it with a new one.

In the current implementation, the layer comprises three instance groups by the service they offer:

- localization group,
- map group,
- path planning group.

Every group provides a REpresentational State Transfer (REST) interface serving as endpoint implemented with the Flask framework exercising Apache2 servers to cater its abilities to re-requesting instances.

The localization group facilitates position estimation by evaluating the collected RFID tag number, by forming a dynamic query to the relational database storing positions of respective tags. The second, map group is accountable for the map management, meaning by interaction with the database, it alters occupancy state for specific cells on obstacle detection by vehicle's distance sensors. Alternatively, vehicles are permitted to utilize this service to reserve traversal path by altering the occupancy status, hence fabricating a collision avoidance functionality.

The path planning group expedites path plan creation for the vehicle, therefore is fundamental for the purpose of the system. The proposed service renders the most complex of the three due to the D\* algorithm utilization with the ability to re-plan the initially calculated path if a vehicle observes a previously unknown hindrance. Distinctively from the previous, the disputed service does not engage in immediate database interaction. Nevertheless, it appertains reliance on the vehicles acquired map representation presented to the service upon request receipt, ensuring the service separability by omitting its reliance on other services.

#### IV. EXPERIMENTAL EVALUATION

The system was evaluated in several experiments to evaluate its functionality. Tests were carried out in environment represented by map of size 10x10 cells with known and unknown obstructions. Furthermore, the maximal number of vehicles operating concurrently was limited to 4. In the first experiment, depicted in the figure 3, a single vehicle was tasked to traverse the environment diagonally. The left-side of the figure depicts the trajectory taken when the obstacle presence was not previously known to the system. Therefore, the vehicle initially followed the direct diagonal path, however on the obstacle detection by the distance sensor, it requested path re-planning resulting in hindrance avoidance by deviating from the diagonal path. The right portion of the picture depicts the trajectory when the obstacle presence was previously acknowledged, therefore the system was able to generate path immediately avoiding the obstruction.

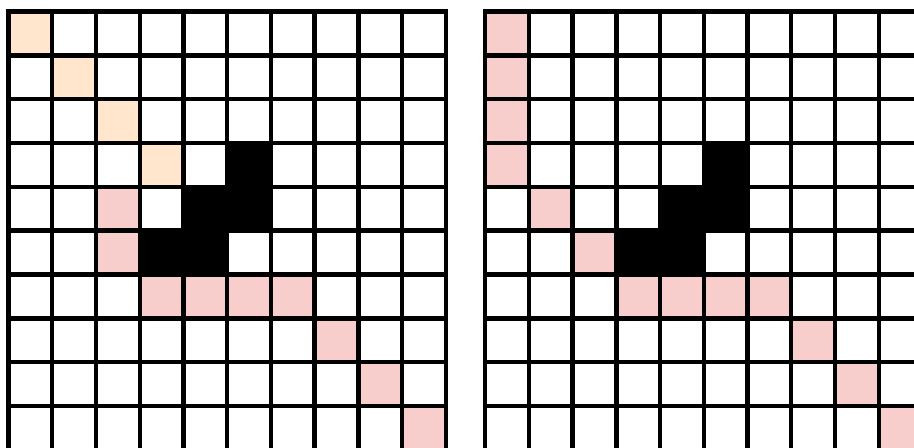


Fig. 3 Path plan generation with known and unknown obstacles.

The second experiment depicted in the figure 4 assessed the system's behaviour during the operation of multiple vehicles, testing the collision avoidance mechanism. The setup required vehicles to traverse to location opposite to their starting point. Primarily, the orange vehicle requested a path plan, resulting in reserving the cells, deferring traversal of the green and red vehicles due to the collision in trajectories. Contrastly, the system generated path for the blue vehicle considering its path to the goal was not obstructed. Following to the traversal of both orange and blue vehicles, the remaining plans were generated.

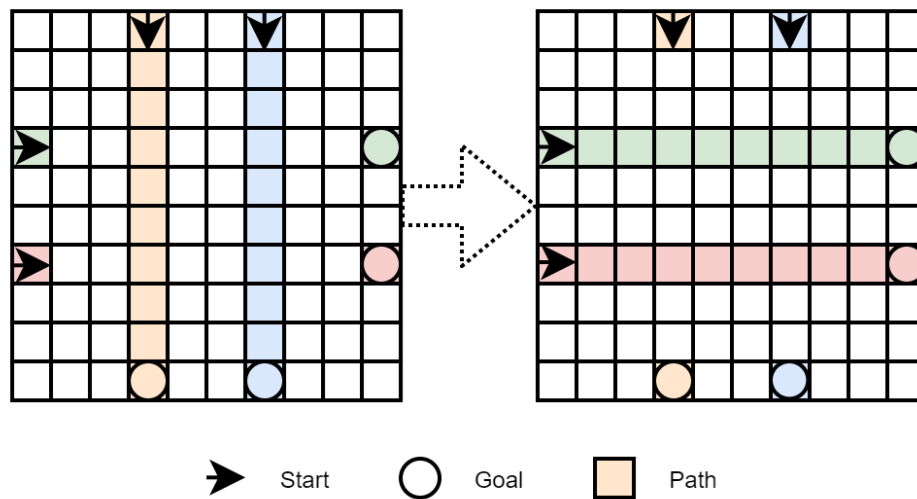


Fig. 4 Path plan generation in presence of multiple vehicles.

## V. CONCLUSION

This paper was concerned with creation of cloud robotics system for autonomous navigation of vehicles. The fundamental problem was management of three tasks, that is localization, mapping and path plan generation for concurrently operating vehicles in an unknown, dynamic environment. To solve those problems, we have created a layered system based on the cloud computing platform enabling data sharing and on-board computation movement to remote servers. The implemented solution was tested in two experiments evaluating the system's performance in each task demonstrating the capability of operation under the defined stipulations.

The designed solution permits an effective operation of multiple vehicles in an unknown environment, managing mapping, localization and path plan generation. The solution ensures scalability and availability of its services. Additionally, the system is not constrained by the computational performance of specific robot unit due to the utilization of cloud computing. Lastly, operation of varying number of vehicles is supported. The subsequent advancement of this solution should focus on implementation of concepts of edge, respectively fog computing to diminish the reliance on stable internet connection and reducing the latency resulting from the remote algorithms execution.

## REFERENCES

- [1] D. Filliat and J.-A. Meyer, "Map-based navigation in mobile robots: I. a review of localization strategies," *Cognitive Systems Research*, vol. 4, no. 4, pp. 243–282, 2003.
- [2] Y. Liu, H. Zhang, and C. Huang, "A novel rgb-d slam algorithm based on cloud robotics," *Sensors*, vol. 19, no. 23, p. 5288, 2019.
- [3] L. Riazuelo, J. Civera, and J. M. Montiel, "C2tam: A cloud framework for cooperative tracking and mapping," *Robotics and Autonomous Systems*, vol. 62, no. 4, pp. 401–413, 2014.
- [4] E. Cardarelli, V. Digani, L. Sabattini, C. Secchi, and C. Fantuzzi, "Cooperative cloud robotics architecture for the coordination of multi-agv systems in industrial warehouses," *Mechatronics*, vol. 45, pp. 1–13, 2017.
- [5] S. Huang and G. Dissanayake, "A critique of current developments in simultaneous localization and mapping," *International Journal of Advanced Robotic Systems*, vol. 13, no. 5, p. 1729881416669482, 2016.
- [6] Y.-L. Lai and J. Cheng, "A cloud-storage rfid location tracking system," *IEEE transactions on magnetics*, vol. 50, no. 7, pp. 1–4, 2014.
- [7] F. Q. Ansari, J. K. Pal, J. Shukla, G. C. Nandi, and P. Chakraborty, "A cloud based robot localization technique," in *International Conference on Contemporary Computing*. Springer, 2012, pp. 347–357.
- [8] F. Duchoň, A. Babinec, M. Kajan, P. Beňo, M. Florek, T. Fico, and L. Jurišica, "Path planning with modified a star algorithm for a mobile robot," *Procedia Engineering*, vol. 96, pp. 59–69, 2014.
- [9] M.-L. Lam and K.-Y. Lam, "Path planning as a service ppaas: Cloud-based robotic path planning," in *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*. IEEE, 2014, pp. 1839–1844.
- [10] G. Mohanarajah, D. Hunziker, R. D'Andrea, and M. Waibel, "Rapyuta: A cloud robotics platform," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 2, pp. 481–493, 2014.
- [11] B. Sebastian and P. Ben-Tzvi, "Physics based path planning for autonomous tracked vehicle in challenging terrain," *J. Intell. Robotic Syst.*, vol. 95, no. 2, pp. 511–526, 2019.

# Radiofrequency Shielding Enhancement of Building Materials by Carbon-Based Coating

*Miloš ŠÁRPATAKY, Iraida KOLCUNOVÁ, Ján ZBOJOVSKÝ, Juraj  
KURIMSKÝ, Peter HAVRAN*

Department of Electric Power Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 040 01 Košice, Slovak republic

milos.sarpataky@tuke.sk, iraida.kolcunova@tuke.sk, jan.zbojovsky@tuke.sk,  
juraj.kurimsky@tuke.sk, peter.havran@tuke.sk

**Abstract** — Regardless of their size, the municipalities are full of artificial electromagnetic smog. In the present-day era of technologies, EM field pollution is increasing rapidly, which can have unfavourable consequences on human health or the correct operation of electronic devices. Just because of that, shielding of EM field is investigated more often. Shielding effectiveness is not only important factor, but reflection and absorption can show properties of different materials from a different perspective. Moreover, the reflection of the EM field can be undesirable in some specific situations because of the accumulation of the EM field in a restricted area. To prevent undesirable interior environment, building materials can reflect or even absorb electromagnetic (EM) waves. The reflection and absorption can be relatively easily improved/modified by a specific coating applied on the material surface. This paper deals with the radiofrequency (RF) electromagnetic interference (EMI) shielding effectiveness (SE) of selected building materials. The measured materials were polystyrene, calcium silicate board, and aerated concrete block. These materials except for polystyrene can be used for blocking EM waves in the frequency range from 3 GHz to 9 GHz. The laboratory tests revealed their relatively high shielding effectiveness, up to 46 dB. However, various amounts of absorbed and reflected RF waves have been measured. After the application of the shielding coat, the shielding effectiveness increased, however, the RF waves were more reflected than absorbed in the material. Shielding coat effectiveness increased approximately by 20 dB within the measured range.

**Keywords** — electromagnetic field, building materials, shielding effectiveness, shielding coating

## I. INTRODUCTION

Nowadays, in the era of wireless technologies global spreading, serious questions arise about their impact on current biological and technical systems or their mutual coexistence. Generally, these technologies radiate electric, magnetic, or electromagnetic (EM) fields into their surroundings, which occur in gradually higher levels of artificial EM radiation like never before in the world. Unwanted EM radiation is being considered as EM artificial pollution.

The challenges associated with the consequences of the increase in EM artificial pollution have led to intensive research efforts around the world. For example, the EMF-Portal is an extensive literature database with an inventory of 31966 publications and 6790 summaries of individual scientific studies on the effects of electromagnetic fields [1]. This open-access database is operated by the Research Center for Bioelectromagnetic Interaction, part of the Institute of Occupational Medicine, RWTH Aachen University, Germany, and the most comprehensive scientific literature database on biological and health-related effects of non-ionizing electromagnetic radiation (frequency range 0-300 GHz) [2]. The database covers articles mainly from Medline / Pubmed and IEEE Explore which are selected by a specific method. The scientific board selects them to the database content after a review of more than four in average relevant articles each day during the previous five years [1]. There is no doubt that it is a frequently examined field of study which deals with electric, magnetic, or EM fields and their impact on cells [3] and [4], cancer or other diseases like Alzheimer's disease [5], [6], [7] and [8], or experiments conducted with EM impact on animals or human being or particular biological systems [9], [10], [11], [12] and [13]. These facts confirm the need to examine protection against the EM field.

Curiously, sometimes the research results are mutually conflicting [14]. Some studies have suggested certain associations between the use of mobile phones and risk of tumors [15], [16] and [15], on the other hand, there are another ones that have reported no significant association between phones and tumors [17], [18] and [19]. However, it is necessary to be careful and examine possible protection in advance. In addition to protective technical equipment, there are global and local standards and regulations. For example, in Slovakia, there is a regulation regarding public health, declared by the Ministry of health of the Slovak Republic [20]. This regulation deals with the exposure and its limit for public in the context of EM radiation in the environment and it is divided by frequency range and the type of examined field. Usually such regulations commonly refer to the authority of the International Commission on Non-Ionizing Radiation Protection (ICNIRP) In general, the limit values for permitted EM non-ionizing radiation are not the same in countries. [21].

Basic principles of protection against unwanted RF radiation include improvement of the shielding around radiation source, personal shielding equipment, protection by distance, and finally, shielding of the restricted area. Data from several studies that have examined the shielding effectiveness of building materials showed various results. Low-cost ceramic bricks with mill scale admixture show the shielding effectiveness ( $SE$ ) around 5 dB (the frequency range from 8 to 9 GHz) [22]. In another work, the ultra-high-performance cement mortar has reported  $SE$  over 25 dB (the frequency range from 1 to 1.2 GHz) [23]. Such works are rare. In general, a relative lack of information on the research of building materials in terms of their RF shielding properties can be recognized today.

The objective of this study is to investigate the possibility of increasing the EM shielding properties of common building materials. The feature modification was done by the application of additional shielding coating. It was followed by investigation of EM field shielding ( $SE$ ) reflection ( $R$ ) and absorption ( $A$ ) in the frequency range (0,9-9 GHz) using the free space measurement method [22].

## II. EXPERIMENTAL METHOD AND MATERIALS

Materials used for this experiment were polystyrene, the calcium silicate board, and aerated concrete block. The test procedure followed the free space comparison method according to the IEEE 299-2006 specification. As reported by this standard  $SE$  is defined as: "The ratio of the signal received (from a transmitter) without the shield, to the signal received inside the shield; the insertion loss when the shield is placed between the transmitting antenna and the receiving antenna" [24]. It means that shielding effectiveness ( $SE_T$ ) is the parameter of particular materials' ability to block the EM energy of a certain frequency when passing through it [25].  $SE_T$  is combination of absorption ( $SE_A$ ), reflection ( $SE_R$ ) and multiple reflection ( $SE_M$ ) [25].

However, due to the growing absorption of waves that are reflected from the internal surface, the multiple reflections can be ignored in case of the shield thickness is greater than the skin depth [26]. If we consider this fact, the parameter  $SE_T$  can be expressed as

$$SE_T = SE_A + SE_R, \quad (1)$$

where  $SE_A$  is absorption and  $SE_R$  is the reflection of EM waves by shielding material both in (dB).  $SE$  can be described also as

$$SE_H = 20 \log \frac{|H_1|}{|H_2|} = 20 \log \frac{|V_1|}{|V_2|}, \quad (2)$$

where  $H_1$  magnetic field measured using the antenna placed in the prescribed configuration in the absence of the enclosure,  $V_1$  is the voltage reading in the absence of the enclosure,  $H_2$  is the magnetic field measured using the antenna placed in the prescribed configuration within the enclosure and  $V_2$  is the voltage reading within the enclosure [24]. For high range  $SE$  (300 MHz to 100 GHz), we use

$$SE_E = 20 \log \frac{|E_1|}{|E_2|}, \quad (3)$$

or

$$SE_P = 10 \log \frac{P_1}{P_2}, \quad (4)$$

where  $E_1$  is the electric field measured using the antenna placed in the prescribed configuration in the absence of the enclosure,  $P_1$  is the power detected in absence of the enclosure,  $E_2$  is the electric field measured using the antenna placed in the prescribed configuration within the enclosure and  $P_2$  is the power detected within the enclosure [24]. If measurements have been made using nonlinear units, the  $SE$  may be directly derived using the following expressions



$$SE_E = |E_1| - |E_2|, \quad (5a)$$

$$SE_H = |H_1| - |H_2|, \quad (5b)$$

$$SE_V = |V_1| - |V_2|, \quad (5c)$$

$$SE_P = P_1 - P_2, \quad (5d)$$

where  $E_1, H_1, V_1$  or  $P_1$  is the reference electric field, magnetic field, voltage, or power intensity measured without the enclosure and  $E_2, H_2, V_2$  or  $P_2$  is the electric field, magnetic field, voltage, or power intensity measured with the enclosure in place [24].

Figure 1 represents the placement of measuring instruments in the high-quality shielded anechoic chamber (Comtest, NL). Agilent N5181A (Agilent, USA) signal generator is connected to the radiating antenna (Fig. 1, A) R&S HF907 (Rohde & Schwarz, USA). It was placed 171 cm from the receiving antenna (Fig. 1, B) Double Ridged Waveguide Horn (RF Spin, CZ). Both antennas were 80 cm above the chamber floor. Measured material was supported in the middle of them (Fig. 1, C). Receiving antenna sent received signal to signal analyzer Agilent N9038A MXE EMI Receiver (Agilent, USA). Reflections and echoes were eliminated in the RF absorbers ((Fig. 1, D). The received signal was processed in the connected computer. The measurement frequency range was from 0.9 to 9 GHz with step 0.1 GHz. Measurement of reflection was made by the same method, but the distribution of the antennas was different. The antennas were on the same side of the chamber and aimed at the measured object at a certain angle. The angle was set approximately as it was a comparison method, which means that the resulting value was calculated as the difference of two values (with shielding material and without shielding material) with the same arrangement of antennas. The antennas were at a height of 80 cm above the chamber floor, and their distance from the measured object was 1 m. The geometric centres of the antennas were 70 cm apart. During the measurement, the position of the antennas did not change, only the shielding material changed.

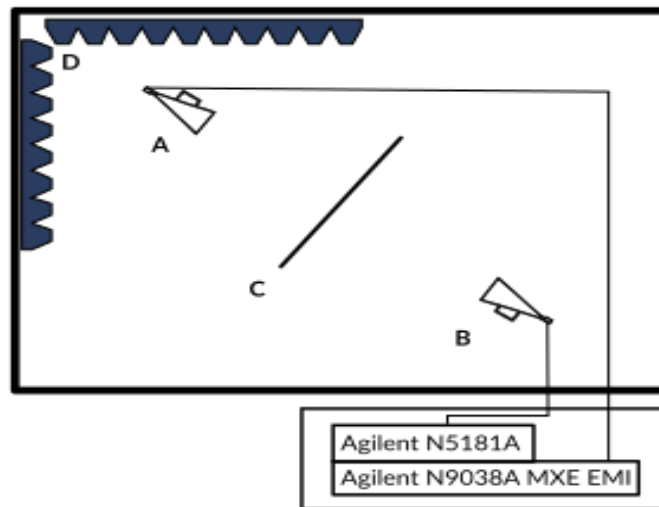


Fig. 1 Schematic presentation of SE measurement in the anechoic chamber. A, B – antennas, C – test plane, D – RF absorbers

SE and reflection ( $R$ ) were calculated using Eq. (5d) from measurement results and absorption ( $A$ ) in (dB) was derived from Eq. (1). Derived equation is

$$A = SE - R. \quad (6)$$

Shielding coat based on carbon fibers was slapped on hardboard (just “shielding coat” below in the text) and attached to a building material closer to radiation and subsequently to receiving antenna so it was possible to compare  $SE$  according to the position of shielding coat.

The first measured building material was expanded polystyrene (EPS) which is characterized as light and strong organic foam that is widely used in European construction, especially as thermal insulation. The thickness of samples was 18 and 30 cm. Aerated concrete block is used for non-load-bearing walls. It has high fire resistance, compressive strength, and low bulk density. The last sample was calcium silicate board used as high-strength and also as acoustic masonry. It is characterized by high heat accumulation and good sound attenuation. In particular, acoustic

partition walls are used for load-bearing and reinforcing walls with high load-bearing capacity.

### III. RESULTS AND DISCUSSION

The following figures (Fig. 2–8) show the measurement results of the shielding effectiveness, reflection, and absorption of building materials with and also without shielding coat, which was measured by the above-described methods.

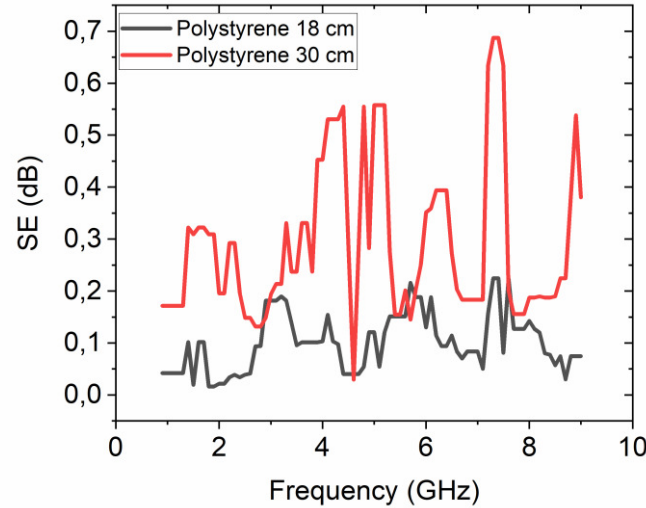


Fig. 2 Shielding effectiveness of two polystyrene desks depending on the frequency.

Figure 2 shows the *SE* of polystyrene of two different thicknesses. Both had almost the same values of *SE* between 0-1 dB, thus we can state that polystyrene as an EMI shielding material is inefficient. In consideration of results, we do not conduct the examination of the reflection and absorption of polystyrene desk because of its negligible *SE* in the raw material.

Next figures represent *SE* of calcium silicate board (CSB), aerated concrete block (ACB) with shielding coat in front of the building material (F), and behind the material (B).

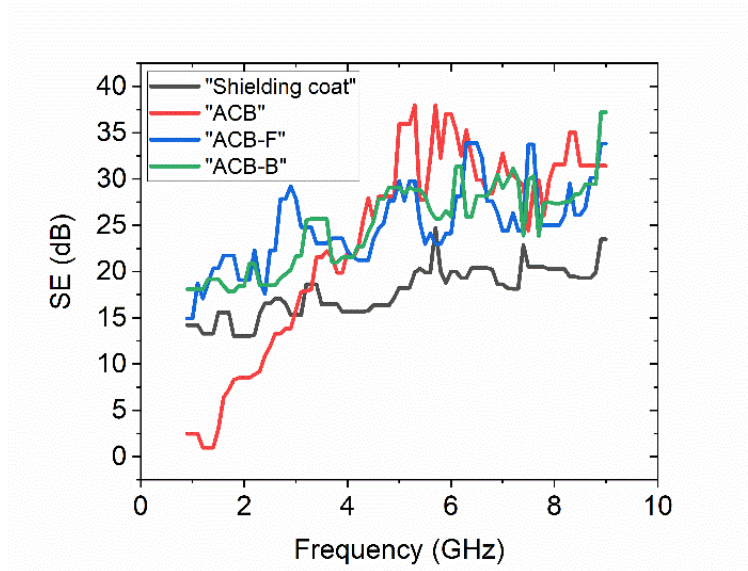


Fig. 3 Graph describing the development of the SE (dB) of the aerated concrete block depending on the frequency (GHz).

In Figure 3, the comparison of the different variants of a position of shielding coat according to the radiation antenna. Comparing the measurement results of the shielding coat and ACB, the difference of SE depends on the frequency. At low frequencies up to 3 GHz, a shielding coat has better shielding properties. Further up to 4 GHz, the shielding effectiveness of the building material improves and begins to exceed the shielding quality of the coat, and from 4 GHz, AC has a greater attenuation than the shielding coat itself. After adding the coat to ACB shielding

properties get better mainly in frequencies under 3 GHz. The position of the shielding coat did not make a significant difference between the results of SE.

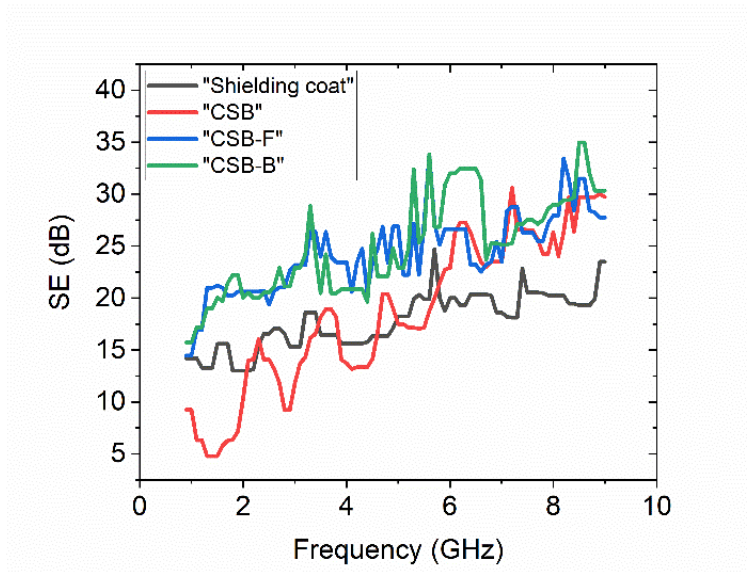


Fig. 4 Graph describing the development of the SE (dB) of calcium silicate board depending on the frequency (GHz).

To identify the difference between Fig. 3 and Fig. 4, ACB had better results than CSB, but curves of both had almost the same course. The difference between the position of shielding coat is not that significant same as it was in Fig.3. Shielding coat improves SE especially in lower frequencies up to 3GHz.

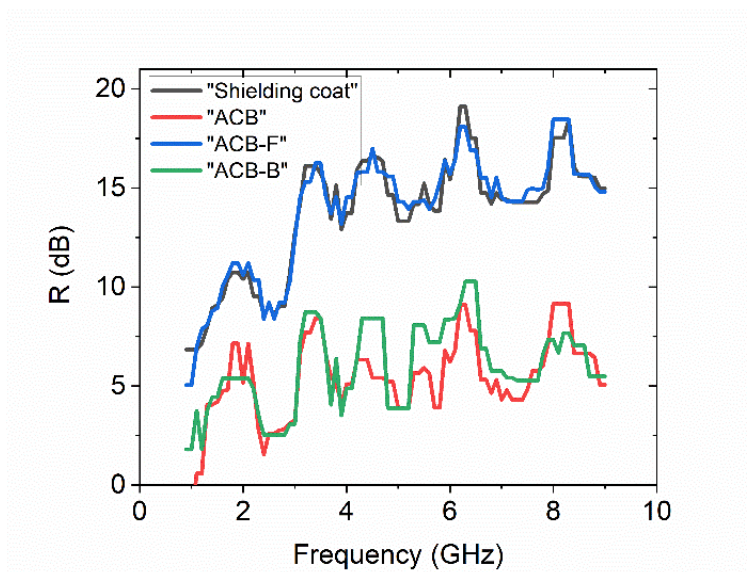


Fig. 5 Graph describing the development of the reflection (dB) of the aerated concrete block depending on the frequency (GHz).

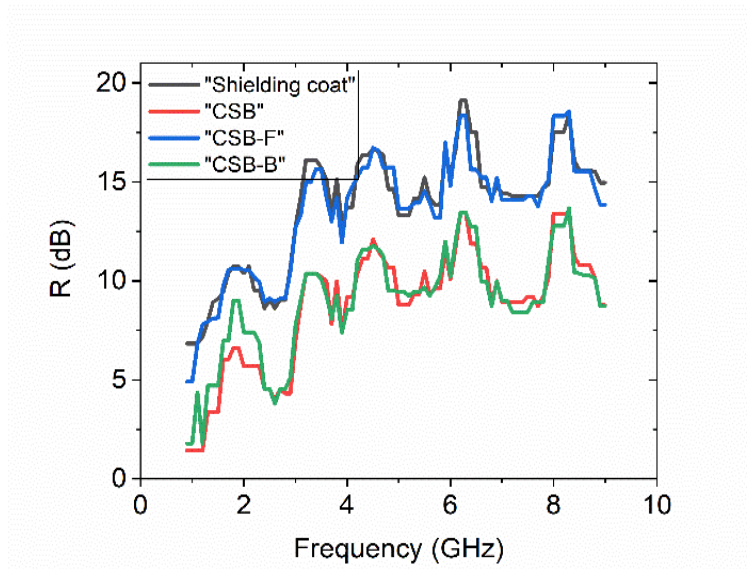


Fig. 6 Graph describing the development of the reflection (dB) of calcium silicate board depending on the frequency (GHz).

Reflection of measured building materials is in Fig. 5 and 6. In both graphs, we have a pair of curves that have almost the same course, because the reflection of connected materials depends on which material is closer to the radiation antenna. If EM waves landed on the shielding coat and the building material is further from radiation antenna values are high and vice versa. If we compare measured materials shielding coat reflects most of the EM waves landing on the material and CSB had higher values of reflection than ACB.

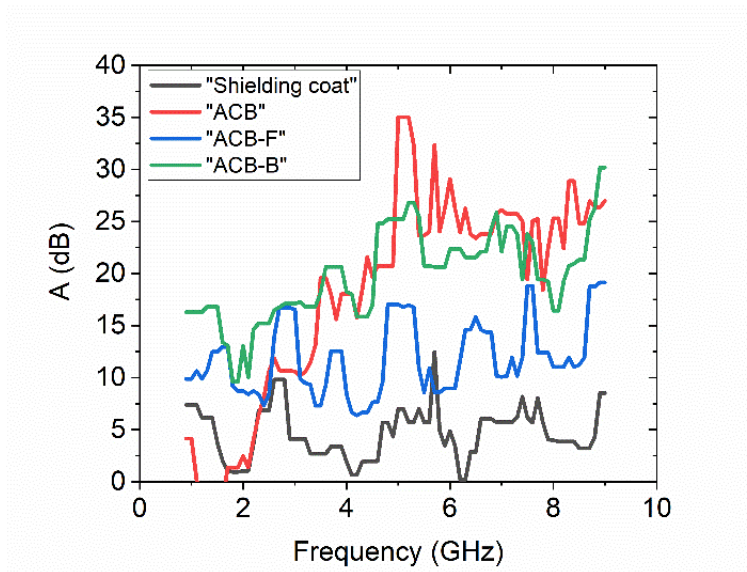


Fig. 7 Graph describing the development of the absorption (dB) of the aerated concrete block depending on the frequency (GHz).

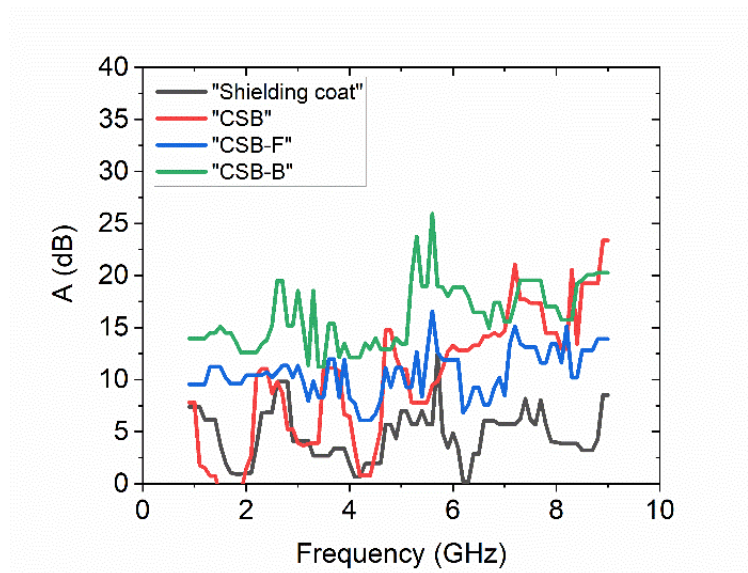


Fig. 8 Graph describing the development of the absorption (dB) of calcium silicate board depending on the frequency (GHz).

Absorption was calculated according to Eq. 6 and results are in Fig. 7 and 8. ACB had the best absorption properties and at some frequencies, this building material absorbs almost all EM waves for example at frequency 5,8 GHz from *SE* 46.65 dB was absorption 42.92 dB so this material absorbs 92 % of landed EM waves. On the other side shielding coat and CSB had zero absorption at some frequencies so they reflect all EM waves landed on their surface.

Table 1 shows *SE* of building materials and their combinations with shielding coat at some specific frequencies frequently used in everyday life. Frequencies used for GPS are around 1.6 GHz and the best results at this frequency had a shielding coat. EM field radiated from mobile phones (UTMS) and Wi-Fi (2.4 GHz) is mostly inhibited by a combination of building material and shielding coat. At higher frequencies of Wi-Fi (5.2 GHz) is the best option ACB with *SE* over 40 dB.

Table 1  
 SE comparison at specific frequencies

Sample	SE (dB)			
	GPS 1,6 GHz	UTMS 2,1 GHz	Wi-Fi 2,4 GHz	Wi-Fi 5,2 GHz
Shielding coating	26.46	13.14	15.42	20.33
ACB	6.4	9.21	11.86	40.93
ACB-F	21.72	19.08	22.3	31.05
ACB-B	17.82	20.97	14.43	32.16
CSB	6.28	16.05	19.34	15.72
CSB-F	18.3	20.62	17.72	22.02
CS-B	18.05	20.01	17.19	22.9

#### IV. CONCLUSION

Building materials were examined by comparison method according to IEEE 299-2006 and measurement showed their different properties. Calcium silicate board (CSB) and aerated concrete block (ACB) had impressive results and we can state that they are possible candidates for shielding the EM field at higher frequencies. By the addition of a shielding coat, it is possible to improve their *SE* primarily at lower frequencies from 0.9 to 3 GHz. Devices used in everyday life mostly radiate EM waves in frequencies below 3 GHz so shielding coat has its meaning in some applications. The biggest disadvantage of shielding coat is its high value of reflection. Reflection of the EM field can cause the accumulation of EM field in some restricted area what can be undesirable. On the other hand, at higher frequencies for example 5.2 GHz, which is the

frequency of Wi-Fi used for shorter distances with higher data rates, the best option is ACB with SE 40.93 dB and its high absorption 37.04 dB at this frequency. To sum up it is possible that experiments with building materials and its shielding efficiency will be examined more often because of the present era of technologies and increasing EM pollution.

#### ACKNOWLEDGMENT

The authors also wish to acknowledge Scientific Grant Agency of the Ministry of Slovak Republic and Slovak Academy of Science for funding experimental works in the frame of project 1/0435/19 and 2/0011/20 and Slovak Research and Development Agency by contract APVV 17-0372, APVV-18-0160 and APVV-15-0438.

#### REFERENCES

- [1] EMF-Portal | Home Available at: <https://www.emf-portal.org/en>
- [2] WHO | EMF research databases. WHO Available at: <http://www.who.int/peh-emf/research/database/en/>
- [3] LEE, Hyung Chul, Mi-Na HONG, Seung Hee JUNG, et al. Effect of extremely low frequency magnetic fields on cell proliferation and gene expression. *Bioelectromagnetics* [online]. 2015, 36(7), 506–516. ISSN 1521-186X. doi:10.1002/bem.21932
- [4] IATCHEVA, Iлона and Ilonka SAYKOVA. FEM Modelling of Heat Effects in Cellular Materials under the Application of Electric Fields. In: 2019 19th International Symposium on Electromagnetic Fields in Mechatronics, Electrical and Electronic Engineering (ISEF):[online]. Nancy, France: IEEE, 2019, s. 1–2 ISBN 978-1-72811-560-3. doi:10.1109/ISEF45929.2019.9097075
- [5] ZUO, Hongyan, Xiao LIU, Yang LI, et al. The Mitochondria/Caspase-Dependent Apoptotic Pathway Plays a Role in the Positive Effects of a Power frequency electromagnetic field on Alzheimer's Disease Neuronal Model: Running title: The effects & mechanism of PF-MF on an AD neuronal model. *Journal of Chemical Neuroanatomy* [online]. 2020, 101857. ISSN 0891-0618. doi:10.1016/j.jchemneu.2020.101857
- [6] ZHANG, Yemao, Jinsheng LAI, Guoran RUAN, et al. Meta-analysis of extremely low frequency electromagnetic fields and cancer risk: a pooled analysis of epidemiologic studies. *Environment International* [online]. 2016, 88, 36–43. ISSN 0160-4120. doi:10.1016/j.envint.2015.12.012
- [7] GRUNDY, Anne, Shelley A. HARRIS, Paul A. DEMERS, et al. Occupational exposure to magnetic fields and breast cancer among Canadian men. *Cancer Medicine* [online]. 2016, 5(3), 586–596. ISSN 2045-7634. doi:10.1002/cam4.581
- [8] RAHMAN, Nur Adilah Abd, Muhammad Mahadi Abdul JAMIL, et al. Pulse Electric Field Effect on Breast Cancer Cell Properties. In: 2019 IEEE 10th Control and System Graduate Research Colloquium (ICSGRC): [online]. 2019, s. 100–103. doi:10.1109/ICSGRC.2019.8837055
- [9] TASIĆ, Tatjana, Maja LOZIĆ, Sofija GLUMAC, et al. Static magnetic field on behavior, hematological parameters and organ damage in spontaneously hypertensive rats. *Ecotoxicology and Environmental Safety* [online]. 2021, 207, 111085. ISSN 0147-6513. doi:10.1016/j.ecoenv.2020.111085
- [10] RUBTSOVA, N. B., S. Yu. PEROV and O. V. BELAYA. The Model of Equivalent Radiofrequency Electromagnetic Field Exposure for Biological Effect Assessment. In: 2019 Photonics Electromagnetics Research Symposium - Spring (PIERS-Spring): 2019 [online]. 2019, s. 1413–1415. ISSN 1559-9450. doi:10.1109/PIERS-Spring46901.2019.9017397
- [11] TETTAMANTI, Giorgio, Anssi AUVINEN, Torbjörn ÅKERSTEDT, et al. Long-term effect of mobile phone use on sleep quality: Results from the cohort study of mobile phone use and health (COSMOS). *Environment International* [online]. 2020, 140, 105687. ISSN 0160-4120. doi:10.1016/j.envint.2020.105687
- [12] KHAKI-KHATIBI, Fatemeh, Alireza NOURAZARIAN, Fatemeh AHMADI, et al. Relationship between the use of electronic devices and susceptibility to multiple sclerosis. *Cognitive Neurodynamics* [online]. 2019, 13(3), 287–292. ISSN 1871-4099. doi:10.1007/s11571-019-09524-1
- [13] RUBTSOVA, N. B., S. Yu. PEROV, O. V. BELAYA, et al. Biological Effects of Power Frequency Electric Field Shielding. In: 2019 Photonics Electromagnetics Research Symposium - Spring (PIERS-Spring): [online]. 2019, s. 1445–1448. ISSN 1559-9450. doi:10.1109/PIERS-Spring46901.2019.9017713
- [14] MYUNG, Seung-Kwon, Woong JU, Diana D. MCDONNELL, et al. Mobile Phone Use and Risk of Tumors: A Meta-Analysis. *Journal of Clinical Oncology* [online]. 2009, 27(33), 5565–5572. ISSN 0732-183X. doi:10.1200/JCO.2008.21.6366
- [15] SADETZKI, S., A. CHETRIT, A. JARUS-HAKAK, et al. Cellular Phone Use and Risk of Benign and Malignant Parotid Gland Tumors--A Nationwide Case-Control Study. *American Journal of Epidemiology* [online]. 2008, 167(4), 457–467. ISSN 0002-9262, 1476-6256. doi:10.1093/aje/kwm325
- [16] LAHKOLA, Anna, Anssi AUVINEN, Jani RAITANEN, et al. Mobile phone use and risk of glioma in 5 North European countries. *International Journal of Cancer* [online]. 2007, 120(8), 1769–1775. ISSN 00207136. doi:10.1002/ijc.22503
- [17] LAHKOLA, A, T SALMINEN, J RAITANEN, et al. Meningioma and mobile phone use--a collaborative case-control study in five North European countries. *International Journal of Epidemiology* [online]. 2008, 37(6), 1304–1313. ISSN 0300-5771, 1464-3685. doi:10.1093/ije/dyn155
- [18] TAKEBAYASHI, T, S AKIBA, Y KIKUCHI, et al. Mobile phone use and acoustic neuroma risk in Japan. *Occupational and Environmental Medicine* [online]. 2006, 63(12), 802–807. ISSN 1351-0711. Available at: doi:10.1136/oem.2006.028308
- [19] SCHÜZ, Joachim, Eva BÖHLER, Gabriele BERG, et al. Cellular Phones, Cordless Phones, and the Risks of Glioma and Meningioma (Interphone Study Group, Germany). *American Journal of Epidemiology* [online]. 2006, 163(6), 512–520. ISSN 1476-6256, 0002-9262. doi:10.1093/aje/kwj068G. 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.
- [20] Slovenská legislatíva | Vyhláška č. 534/2007 Z. z. o podrobnostiach o požiadavkách na zdroje elektromagnetického žiarenia a na limity expozície obyvateľov elektromagnetickému žiareniu v životnom prostredí Available at: <https://www.law-reg.org/legislativa/predpisy/vyhlaska-c-5342007-z-z-o-podrobnostiach-o-poziadavkach-na-zdroje-elektromagnetickeho-ziarenia-a-na-limity-expozicie-obyvatelov-elektromagnetickemu-ziarieniu-v-zivotnom-prostredii>

- [21] EMF-Portal | Limit values compared internationally Available at: <https://www.emf-portal.org/en/cms/page/home/more/limits/limit-values-compared-internationally>
- [22] BANTSIS, G., C. SIKALIDIS, M. BETSIOU, et al. Electromagnetic absorption, reflection and interference shielding in X-band frequency range of low cost ceramic building bricks and sandwich type ceramic tiles using mill scale waste as an admixture. *Ceramics International* [online]. 2011, 37(8), 3535–3545. ISSN 0272-8842. doi:10.1016/j.ceramint.2011.06.010
- [23] YOO, Doo-Yeol, Min-Chang KANG, Hong-Joon CHOI, et al. Electromagnetic interference shielding of multi-cracked high-performance fiber-reinforced cement composites – Effects of matrix strength and carbon fiber. *Construction and Building Materials* [online]. 2020, 261, 119949. ISSN 0950-0618. doi:10.1016/j.conbuildmat.2020.119949
- [24] IEEE 299-2006 - IEEE Standard Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosure Available at: <https://standards.ieee.org/standard/299-2006.html>
- [25] SHUKLA, Vineeta. Review of electromagnetic interference shielding materials fabricated by iron ingredients. *Nanoscale Advances* [online]. 2019, 1(5), 1640–1671. ISSN 2516-0230. doi:10.1039/C9NA00108E
- [26] BHATTACHARJEE, Yudhajit, Dipanwita CHATTERJEE and Suryasarathi BOSE. Core–Multishell Heterostructure with Excellent Heat Dissipation for Electromagnetic Interference Shielding. *ACS Applied Materials & Interfaces* [online]. 2018, 10(36), 30762–30773. ISSN 19448244, 19448252. doi:10.1021/acsami.8b10819

# Vybrané metódy bezsnímačového riadenia synchronného motora s permanentnými magnetmi

Viktor Petro

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

viktor.petro@tuke.sk

**Abstrakt** – V tomto článku sú uvedené bežne používané prístupy bezsnímačového riadenia synchronného motora s permanentnými magnetmi (SMPM). Ako prvé sú opísané a zatriedené metódy využívajúce injektáž napätia/prúdu. Ozvy týchto injektovaných signálov nesú informáciu o aktuálnej polohe rotora. V ďalšej časti článku sú uvedené pozorovatele využívajúce indukované napätie a model stroja. Z oboch prístupov je vybratá jedná metóda, ktorá je detailnejšie opísaná.

**Kľúčové slová** – Synchronný motor s permanentnými magnetmi, bezsnímačové riadenie, injektáž vysokofrekvenčného signálu, pozorovateľ, kľzavý režim

## I. ÚVOD

Synchronné motory s permanentnými magnetmi (SMPM) sú často používané v priemyselných aplikáciách, ako aj v zariadeniach určených pre domácnosti, kvôli vysokej účinnosti, dynamickým charakteristikám a riaditeľnosti. Najčastejšie používaná stratégia riadenia SMPM je vektorové riadenie. Zjednodušený princíp vektorového riadenia je rozloženie statorového prúdu na momentotvornú a tokotvornú zložku. Následne je riadenie komplexného trojfázového striedavého motora založené na rovnakom princípe, ako to je pri jednosmernom cudzobudenom motore. V prípade jednosmerného cudzobudeného stroja, prúd v rotore vytvára moment a prúd v budiacom vinutí vytvára tok stroja. Pretože SMPM má permanentné budenie vytvorené permanentnými magnetmi, budiaci prúd je okrem prípadu odbudzovania nastavený na nulu.

Vo vektorovom riadení sa používa transformácia zo statorového do rotorového súradnicového systému, ktorým sa docieli odväzbenie momentotvornej a tokotvornej zložky prúdu. Táto transformácia sa označuje ako Parková transformácia, pre ktorú je nevyhnutná presná elektrická poloha rotora. Vo všeobecnosti sa využívajú snímače polohy, ktoré sú mechanicky spojené s rotorom. Mechanická poloha je následne prevedená na elektrickú použitím počtom párov pólov stroja. Takéto snímače spoľahlivo fungujú v celom pracovnom rozsahu stroja. Avšak použitie snímača znižuje celkovú robustnosť systému. Okrem toho, eliminácia snímača polohy ma aj mnoho iných prínosov, napr. menšia náročnosť pohonu na montážne rozmery, zvýšenie spoľahlivosti, úspora na cene, atď. Nevýhodou bezsnímačového riadenia je menšia presnosť riadenia a náklady na vývoj softvéru.

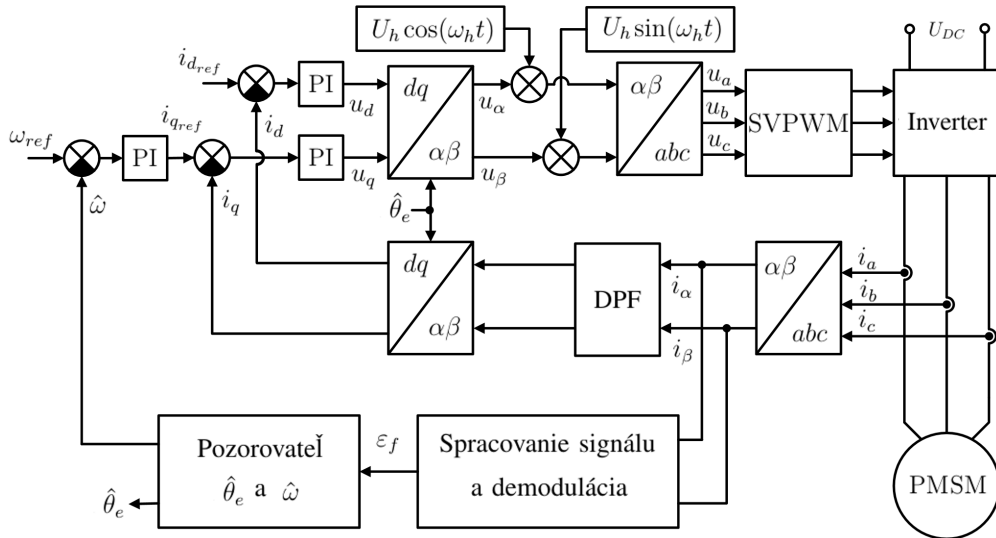
## II. BEZSNÍMAČOVÉ RIADENIE SMPM

Ako prvé je potrebné ujasniť, že bezsnímačovým riadením pohonu je myslené také riadenie, pri ktorom sa nepužíva snímač polohy, resp. rýchlosti. Snímače prúdov sú samozrejme použité a teda namiesto snímania polohy sa využije model motora a elektrické veličiny ako sú prúd, napätie a parametre stroja. Základné prístupy bezsnímačového riadenia môžu byť rozdelené do nasledovných skupín [1]:

- metódy založené na injektáži prúdu/napätia,
- metódy založené na modely motora,
- metódy využívajúce umelú inteligenciu.

Metódy založené na injektáži prúdu/napätia dokážu správne pozorovať polohu pri pomalých a nulových rýchlostiach. Vysokofrekvenčná (VF) injektáž signálu je viac preferovaná ako nízko-frekvenčná (NF) injektáž. Signál môže byť injektovaný v statorovom alebo v rotorovom súradnicovom systéme. Ďalej signál môže byť superponovaný na základnú harmonickú, alebo môže





Obr. 1 Bloková schéma bezsnímačového vektorového riadenia použitím metódy injektáže rotačného VF signálu.

byť injektovaný nezávisle v čase, keď vektorové riadenie je prerušené. Preto metódy založené na injektáži prúdu/napätia môžeme ďalej rozdeliť do nasledovných skupín:

- VF metódy založené na injektáži signálu so sínusovým priebehom,
  - injektáž rotačného VF signálu so sínusovým priebehom,
  - injektáž pulzujúceho VF signálu so sínusovým priebehom,
- VF metódy založené na injektáži signálu s obdĺžnikovým priebehom,
- VF metódy založené na injektáži pulzujúceho signálu.

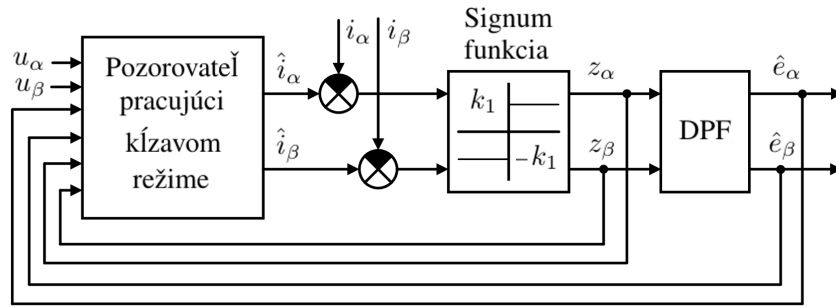
Metódy založené na modeli motora väčšinou využívajú indukované napätie na estimáciu polohy a uhlovej rýchlosti motora. Pretože amplitúda indukovaného napätia je nedostatočne veľká na spoľahlivé pozorovanie polohy pri nízkych rýchlostiach, tieto metódy sa využívajú pre oblasť nad 5%–10% nominálnej uhlovej rýchlosti [2]. Do tejto skupiny patrí napr. Luenbergerov pozorovateľ, Kalmanov filter, pozorovateľ toku stroja, pozorovateľ pracujúci v kľzavom režime, pozorovateľ využívajúci adaptačný model motora, atď.

### III. INJEKTÁŽ ROTAČNÉHO VF SIGNÁLU SO SÍNUSOVÝM PRIEBEHOM

Injektáž rotačného VF signálu so sínusovým priebehom prvý krát publikoval R. D. Lorenz [3]. Táto metóda je taktiež označovaná ako  $\alpha\beta$  injektáž, keďže rotačný signál je injektovaný v statorovom súradnicovom systéme. Pre dva injektované napätia, ktoré sú na seba kolmé, majú konštantnú amplitúdu a frekvenciu platí [4]:

$$\begin{bmatrix} u_{\alpha h} \\ u_{\beta h} \end{bmatrix} = U_h \begin{bmatrix} \cos(\omega_h t) \\ \sin(\omega_h t) \end{bmatrix}, \quad (1)$$

kde  $h$  označuje že ide o vysokofrekvenčný signál,  $U_h$  je amplitúda a  $\omega_h$  je frekvencia injektovaného VF signálu. Bloková schéma bezsnímačového riadenia SMPM použitím metódy injektáže rotačného VF signálu so sínusovým priebehom je zobrazená na Obr. 1. Superponovaný VF napäťový signál vyvolá v statorovom vinutí VF prúd, ktorý nesie informáciu o polohe rotora. Zosnímané prúdy sa najprv transformujú do dvojfázovej statorovej sústavy  $\alpha\beta$  a následne sa odfiltruje základná harmonická použitím hornopriepustného filtra, alebo pásmovej priepusti. Následne sa vysokofrekvenčný signál demoduluje, čím sa získa signál  $\varepsilon_f$ . Demodulačný proces je opísaný napr. v [5], [6]. Signál  $\varepsilon_f$  je proporcionálny rozdiel medzi skutočnou polohou rotora  $\theta_e$  a pozorovanou  $\hat{\theta}_e$ . Následne sa použije pozorovateľ, ktorý je založený na Luenbergerovom pozorovateli na estimáciu polohy a uhlovej rýchlosti motora. Pozorovateľ má za úlohu potlačiť odchýlku  $\varepsilon_f$  do nulovej hodnoty a tým zabezpečiť konvergenciu pozorovanej polohy ku skutočnej, t.j.  $\hat{\theta}_e \rightarrow \theta_e$  [3], [7]. Pretože snímaný prúd má aj vysokofrekvenčnú zložku, nemôže byť priamo použitý v procese riadenia. Na odstránenie VF zložky je použitý dolnopriepustný filter (DPF). Pozorovaný elektrický uhol  $\hat{\theta}_e$  je následne použitý v Parkovej transformácii a pozorovaná uhlová rýchlosť predstavuje spätnú väzbu otáčkového regulátora. Hlavná nevýhoda tejto metódy je momentové zvlnenie motora. Okrem toho, použitie dolnopriepustného a hornopriepustného filtra je nevyhnutné v procese pozorovania a regulácie.



Obr. 2 Bloková schéma pozorovateľa pracujúceho v kĺzavom režime.

#### IV. POZOROVATEĽ PRACUJÚCI V KĹZAVOM REŽIME

Pozorovateľ pracujúci v kĺzavom režime je využívaný v priemysle kvôli robustnosti a jednoduhosti. Firma Texas Instruments ponúka knižnicu pre implementáciu pozorovateľa na mikroprocesor opísaný v [8]. Firma Microchip má tiež vyvinuté riešenie, ktoré je možné nájsť v [9]. Firma Whirlpool implementovala do svojich produktov riadenie, využívajúce pozorovateľa pracujúceho v kĺzavom režime [10]. NXP našlo uplatnenie tohto pozorovateľa pri riadení kompresorov [11]. V tomto prípade je použitý pozorovateľ štvrtého rádu, pričom zložky indukovaného napätia sú považované za stavové veličiny. Podobný pozorovateľ štvrtého radu bol použitý aj v [12], doplnený o adaptívny výpočet uhlovej rýchlosti.

V tomto článku bude uvedený pozorovateľ 2. rádu, podobne ako v [4]. V tomto prípade je indukované napätie vnímané ako porucha, ktorá je pomocou pozorovateľa kompenzovaná. Matematický model pozorovateľa je väčšinou vyjadrený v  $\alpha\beta$  statorovom súradnicovom systéme a vyzerá nasledovne:

$$\begin{bmatrix} \dot{\hat{i}}_{\alpha} \\ \dot{\hat{i}}_{\beta} \end{bmatrix} = \frac{1}{L_s} \begin{bmatrix} -R_s & 0 \\ 0 & -R_s \end{bmatrix} \begin{bmatrix} \hat{i}_{\alpha} \\ \hat{i}_{\beta} \end{bmatrix} + \frac{1}{L_s} \begin{bmatrix} u_{\alpha} - \hat{e}_{\alpha} - z_{\alpha} \\ u_{\beta} - \hat{e}_{\beta} - z_{\beta} \end{bmatrix}, \quad (2)$$

kde  $\hat{i}_{\alpha}, \hat{i}_{\beta}$  sú pozorované prúdy,  $u_{\alpha}, u_{\beta}$  sú želané napätia,  $L_s, R_s$  sú indukčnosť a odpor jednej fázy stroja,  $\hat{e}_{\alpha}, \hat{e}_{\beta}$  sú pozorované indukované napätia a  $z_{\alpha}, z_{\beta}$  sa označujú ako kompenzačné spätné väzby pozorovateľa, platí:

$$\begin{bmatrix} z_{\alpha} \\ z_{\beta} \end{bmatrix} = k \begin{bmatrix} \text{sign}(\hat{i}_{\alpha} - i_{\alpha}) \\ \text{sign}(\hat{i}_{\beta} - i_{\beta}) \end{bmatrix}, \quad (3)$$

pričom  $k$  je zosilnenie pozorovateľa a podľa [13] platí:

$$k > \max(|e_{\alpha}|, |e_{\beta}|). \quad (4)$$

Pozorované indukované napätia je možné získať nasledovne:

$$\begin{bmatrix} \hat{e}_{\alpha} \\ \hat{e}_{\beta} \end{bmatrix} = \frac{\omega_c}{s + \omega_c} \begin{bmatrix} z_{\alpha} \\ z_{\beta} \end{bmatrix}, \quad (5)$$

kde  $\omega_c$  reprezentuje frekvenciu útlmu dolnopriepustného filtra. Z pozorovaného indukovaného napätia sa následne získa pozorovaná poloha:

$$\hat{\theta}_e = -\arctan\left(\frac{\hat{e}_{\alpha}}{\hat{e}_{\beta}}\right). \quad (6)$$

Blokový diagram pozorovateľa pracujúceho v kĺzavom režime je zobrazený na Obr. 2.

#### V. ZÁVER

Bezsnímačové riadenie motorov je dlhodobo skúmané z dôvodu možných úspor na cene pohonu. V tomto článku boli zatriedené jednotlivé metódy a detailnejšie opísané vybrané prístupy. Ako to z článku vyplýva, jednotlivé metódy osobitne nepokrývajú celú oblasť rýchlostného riadenia. Metódy založené na injektáži sa využívajú pri nízkych rýchlostiach a metódy založené na indukovanom napätí a modely motora sa používajú od 5%–10% nominálnej uhlovej rýchlosti. Pre plnohodnotné riadenie v celej pracovnej oblasti sa používajú hybridné riešenia, ktoré spájajú tieto dve metódy [14].

## REFERENCES

- [1] J. Agrawal and S. Bodkhe, "Experimental study of low speed sensorless control of PMSM drive using high frequency signal injection," *Advances in Electrical and Electronic Engineering*, vol. 14, no. 1, pp. 29–39, 2016.
- [2] Z. Ma and R. Kennel, "System-on-Chip sensorless control of PMSM combining signal injection and flux observer," *Conference Proceedings - 2012 IEEE 7th International Power Electronics and Motion Control Conference - ECCE Asia, IPEMC 2012*, vol. 2, pp. 1201–1205, 2012.
- [3] P. L. Jansen and R. D. Lorenz, "Transducerless Position and Velocity Estimation in Induction and Salient AC Machines," *IEEE Transactions on Industry Applications*, vol. 31, no. 2, pp. 240–247, 1995.
- [4] G. Wang, G. Zhang, and D. Xu, *Position sensorless control techniques for permanent magnet synchronous machine drives*, 2019.
- [5] Z. Chen, M. Tomita, S. Doki, and S. Okuma, "An extended electromotive force model for sensorless control of interior permanent-magnet synchronous motors," *IEEE Transactions on Industrial Electronics*, vol. 50, no. 2, pp. 288–295, 2003.
- [6] M. Koteich, A. Messali, and S. Daurette, "Self-sensing control of the externally-excited synchronous machine for electric vehicle traction application," *Conference Proceedings - 2017 8th International Symposium on Sensorless Control for Electrical Drives, SLED 2017*, no. 6, pp. 91–96, 2017.
- [7] K. W. Van, "High-Resolution Velocity Estimation for All-Digital, ac Servo Drives," *IEEE Transactions on Industry Applications*, vol. 27, no. 4, pp. 701–705, 1991.
- [8] T. INSTRUMENTS, "Digital Motor Control Software Library.pdf," no. October, 2003.
- [9] Microchip, "AN1078-Sensorless Field Oriented Control of a PMSM," *an1078*, vol. PP, no. 1, pp. 1–6, 2019.
- [10] S. Chi, Z. Zhang, and L. Xu, "Sliding-mode sensorless control of direct-drive PM synchronous motors for washing machine applications," *IEEE Transactions on Industry Applications*, vol. 45, no. 2, pp. 582–590, 2009.
- [11] Freescale Semiconductor, "Sensorless PMSM Vector Control with a Sliding Mode Observer for Compressors Using MC56F8013 Devices Supported : Home Page :," 2008.
- [12] Z. Chen, M. Tomita, S. Doki, and S. Okuma, "New adaptive sliding observers for position-and velocity-sensorless controls of brushless dc motors," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 3, pp. 582–591, 2000.
- [13] M. Elbuluk and C. Li, "Sliding Mode Observer for Wide-Speed Sensorless Control of PMSM Drives," *Conference Record - IAS Annual Meeting (IEEE Industry Applications Society)*, vol. 1, pp. 480–485, 2003.
- [14] H. Iura, M. Inazumi, T. Kamei, and K. Ide, "Hybrid sensorless control of IPMSM for direct drive applications," *2010 International Power Electronics Conference - ECCE Asia -, IPEC 2010*, pp. 2761–2767, 2010.

# Implementácia RSS algoritmu v procese meranie sieťovej prevádzky

<sup>1</sup>Martin HASIN, <sup>2</sup>Martin HAVRILLA

<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.hasin@tuke.sk, <sup>2</sup>martin.havrilla@tuke.sk

**Abstrakt** – Táto práca sa zaoberá základnými metódami na zber a analýzu toku dát v skúmanej sieti. Úvodom práca rozoberá problematiku monitorovania sietí v čase a taktiež rozoberá narastajúcu potrebu na monitorovanie sietí v súčasnosti. V neskoršej časti sa tento článok zaoberá netflow analyzátormi a definuje topológiu meracieho systému s využitím technológie RSS počas procesu merania. Záver tohto článku sa venuje zhodnoteniu a zhrnutiu merania a nastavuje potrebné zlepšenia pre dosiahnutie kvalitnejších a presnejších výsledkov.

**Kľúčové slová** – NetFlow, NFStream, RSS

## I. ÚVOD

Sieťová prevádzka, ktorú je potrebné monitorovať v čase narastá, čím narastá aj potreba zvyšovania výkonu zariadenia na analýzu dát. Meranie dát spočíva v načítavaní paketov zo živej sieťovej prevádzky, nasleduje analýza a prevod na NetFlow protokol[1]. Pri tomto spôsobe monitorovania je možné pružne reagovať na rastúcu požiadavku na výkon bez zmeny finančne nákladných sieťových komponentov. Na meranie bol zvolený systém NFstream, ktorý dokáže pomocou knižnice v Pythone merať a súčasne vykonávať analýzu sieťovej prevádzky s minimálnym oneskorením. Pri budovaní systému, ktorého oneskorenie má byť čo najmenšie bolo potrebné rozložiť meranie na viacero jadier procesora. Tento prístup zrýchlil meranie a vykonávanie viacerých analytických modelov súčasne bez straty výkonu, alebo zväčšení oneskorenia. Avšak nárastom sieťovej prevádzky, ktorú je potrebné merať sa narazilo na limity výpočtového jadra ktoré malo prepočítavať sieťovú prevádzku smerujúcu zo sieťovej karty. Tento problém je možné riešiť využitím technológie RSS ktorá zabezpečuje rozkladanie sieťovej prevádzky na viacero jadier procesora. Rozložením tejto sieťovej prevádzky získava systém možnosť analyzovať oveľa vyššiu kapacitu sieťovej prevádzky. Rozloženie výkonu takisto urýchľuje analýzu, keďže nedochádza k zbytočnému preťaženiu jedného jadra pričom ostatné sú nevyťažené.

## II. NFSTREAM NETFLOW ANALYZER

Najpoužívanejším protokolom na analýzu siete je NetFlow. Tento protokol vyvinula spoločnosť CISCO[2], primárne ako doplnkovú službu na ich smerovačoch. Hlavným využitím tohto protokolu je monitorovanie sieťovej prevádzky na základe IP sieťových tokov pomocou ktorých je možné získať pohľad do siete v reálnom čase. Počas vývoja vznikli rôzne verzie NetFlow protokolu. Väčšina zo vzniknutých verzií tohto protokolu nebola uvoľnená pre širokú verejnosť a využívala sa hlavne vo vedeckej oblasti. Prvá verzia bola vydaná v roku 1996 a poskytovala iba základné údaje o meranej sieti. Medzi tieto údaje patrili informácie o IP adresách, iba v protokole IPv4. Vývoj tohto protokolu napredoval a terajšia verzia 10 nadobudla podporu ďalších protokolov ako IPv6, MPLS, BGP[3]. Túto verziu je možné v literatúre nájsť pod názvom IPFIX[4].

Framowork NFStream je komponent, ktorý využíva programovacie prostredie Python. Tento framework umožňuje analyzovať dáta, ktoré vstupujú do programu dvoma spôsobmi. Prvý spôsob je využitie PCAP súboru, ktorý bol vyhotovený meraním v skúmanej sieti, prípadne na meranom sieťovom zariadení akým je napríklad server. Tento prístup je však off-line riešenie analýzy siete. Namerané dáta sú vhodné na cvičenie Machine Learning s preddefinovanými vzorkami nežiaducej prevádzky v sieti[5]. Druhý spôsob analýzy je využitie on-line čítania dát z sieťového adaptéra s využitím knižnice libpcap[6]. Na meraní sieťový adaptér sa pomocou technológie Port Mirror

zrkadlí požadovaný sieťový tok z produkčného sieťového komponentu, najčastejšie sa jedná o router, kde prebieha meraná prevádzka. Pomocou NFStream frameworku je možné rýchlo a flexibilne analyzovať dáta v reálnom čase s minimálnym oneskorením. Vznikajúce oneskorenie je možné eliminovať zvýšením výkonu procesora na ktorom prebieha analýza. Táto analýza si vyžaduje optimálne nastavenie parametrov ML a najmä potrebuje disponovať vyšším výkonom procesora na jednotlivé jadrá. Vplyv na výkon má rovnako aj daný sieťový adaptér, ktorý musí obsahovať rozloženie sieťovej prevádzky do viacerých jadier procesora.

Framowork NFStream umožňuje merané dáta spracovávať a analyzovať v reálnom čase. Je možné využiť vlastné pluginy ktorými je možné vytvoriť vlastné atribúty netflow exportu. Framwrok NFStream disponuje nasledovnými funkcionalitami:

- Vysoká rýchlosť - Tento framework je navrhnutý na rýchlu analýzu siete za použitia knižnice libpcap.
- L7 analýza - Framework umožňuje hĺbkovú inšpekciu paketov s využitím nDPI[7]. Tento proces umožňuje rýchlo zistiť druh aplikácie, ktorá komunikuje pomocou metadát (TLS, SSH, DHCP, HTTP ...)
- Štatistické funkcie - Framework umožňuje zaznamenávať štatistické informácie o sieťových tokoch, ktoré sa vypočítavajú až po ukončení spojenia. Najčastejšie štatistiky: sequence of first n packets sizes, inter arrival times and directions
- Flexibilita - framework umožňuje vytvárať vlastné moduly pomocou ktorých je možné doplniť nové potrebné funkcionality v jazyku python.
- Machine Learning - výstupy z frameworku je možné implementovať v Machine Learning modeloch pomocou knižníc v Pythone[8]. Výsledné natrénované modely je možné rýchlo implementovať do reálnej prevádzky pomocou NFPlugin.

Namerané a analyzované dáta je možné následne zasielať do nerelačných databáz, prípadne ukladať vo formáte CSV. Ukladanie dát je riešenie pomocou knižníc v pythone, ktoré umožňujú vytvoriť priame prepojenie do zvolenej databázy. Pri exporte je možné zvoliť aj postup ukladania do CSV, ktoré je možné následne pomocou inej aplikácie posielat a spracovávať v externých databázach.

### III. TOPOLOGIA MERACIEHO SYSTÉMU

Topológia využitá na meranie sieťovej prevádzky pozostávala z samostatného servera, ktorý analyzoval sieťovú prevádzku z hlavného sieťového routra. Meraná sieťová prevádzka bola na server posielaná s využitím technológie SPAN. Táto technológia replikovala sieťovú prevádzku hlavného internetového pripojenia na server, kde prebiehalo meranie.

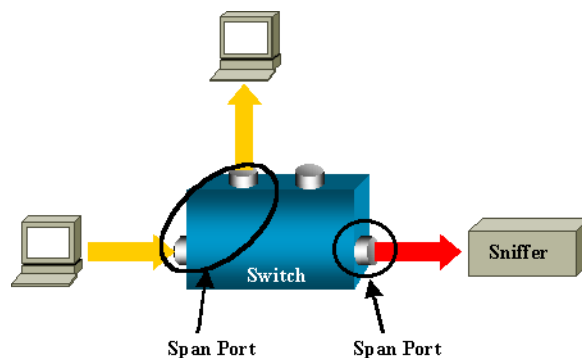


Fig. 1 Topológia merania sieťovej prevádzky v Cisco zariadení

Na Obr. 1 je znázornená topológia zapojenia, ktorá pozostáva zo sieťového prvku Cisco ktorý vytvára destination span port, kde je pripojená sonda vytvárajúca meranie. Výhodou tohto zapojenia je využitie externého servera na vytváranie analýzy, čím sa znižuje zaťaženie sieťového routra a zvyšuje sa flexibilita pri vytváraní merania.

Server na meranie pozostával z HPE DL585 G8 ktorý disponuje štyrmi fyzickými procesormi Intel E5-4640. Týmto sa získalo až 32 fyzických jadier a 64 vlákien, ktoré je možné využiť na spúšťanie analýz. Najdôležitejším bodom bolo získať najväčší počet fyzických jadier, na ktorých dochádza k najrýchlejšiemu prepočtu analýzy. Pri vytváraní NetFlow bol zvolený framework NFStream, ktorý umožňuje rozloženie analýzy na všetky jadra procesora.

Na Obr. 2 je znázornená bloková schéma priebehu merania s využitím NFStream frameworku. Táto schéma reprezentuje jeden blok merania, ktorý beží na jednom jadre procesora, pričom

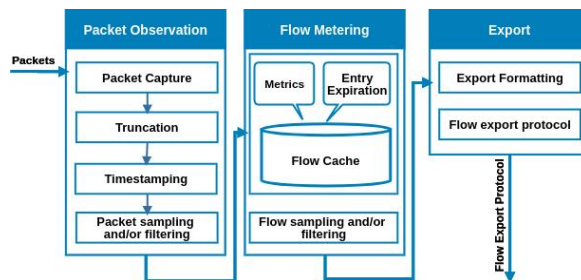


Fig. 2 Topológia NFstream

je možné počet týchto blokov zvoliť manuálne, alebo si systém automaticky zvolí maximálnu hodnotu podľa aktuálneho počtu jadier procesora.

Meranie pozostáva z viacerých krokov, kde prvý krok je rozdelenie prichádzajúcich paketov na viacej blokov o veľkosti 1500 bitov a následne ich rozdelenie medzi skupinu analyzovaných procesov. Analýza v bloku "Flow mettering" umožňuje zbierať údaje o sieťovej prevádzke a štatistické údaje. V tomto kroku sa pakety analyzujú knižnicou NDPi, ktorá dokáže zistiť aký druh komunikačného protokolu sa využíva v komunikácií. Namerané dáta sa ukladajú do Cache v RAM ktorá po jej naplnení sa presúva do bloku Export. Tento blok je pre všetky procesy spoločný kde dochádza k exportovaniu nameraných dát v príslušnom formáte. V tomto kroku je možné dáta finálne upraviť a exportovať napríklad do súboru CSV, prípadne do nerelačných databáz ElasticSearch, alebo iných.

#### IV. VYUŽITIE RSS V PROCESSE MERANIA

Pri meraní na viacjadrových systémoch bolo potrebné využitie Recive side scaling (RSS) [9]. Táto technológia umožňuje efektívne rozkladať sieťovú prevádzku medzi viac jadrové systémy CPU. Posielanie paketov danému jadrú začína už na sieťovej karte, ktorá rozkladá prevádzku pomocou filtra, ktorý daný paket priradí špecifickému jadrú procesora. Pri použití tejto technológie je potrebné voliť viacjadrové procesory, kde využitie vlákien neprinesie zvýšenie výkonu, nakoľko by dochádzalo k využívaniu rovnakého jadra. Technológia umožňuje dve možnosti implementácie Recive side scaling:

- RSS - na hardvérovej báze
- SWRSS na softvérovej báze

Medzi najčastejšie využívanú možnosť sa považuje RSS na hardvérovej báze, ktorá dokáže vykonávať rýchlejšie rozkladanie paketov medzi jadra procesora, ako je to pri využití softvérovej varianty[10]. Softvérová varianta sa využíva hlavne v prípadoch adaptérov, prípadne systémov, ktoré neumožňujú pracovať s hardvérom určeným na túto činnosť.

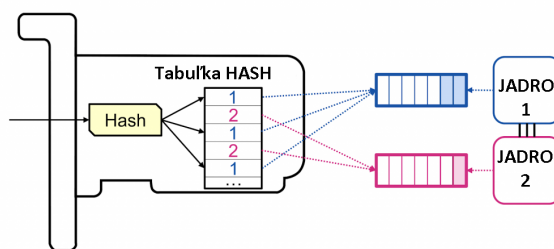


Fig. 3 Rozloženie záťaže pomocou RSS algoritmu

Na Obr.3 je znázornená metóda rozkladania paketov na viacero jadier procesora. Zvýšením počtu jadier na ktorých dochádza k analýze sa zvyšuje aj sieťová priepustnosť meracieho systému. Rozkladanie do viacerých jadier začína už na strane sieťovej karty, kde sa označujú pakety pomocou HASH funkcie. Pri vypnutej technológii RSS dochádza k označovaniu paketov rovnakou HASH funkciou. Takto označený obsah sa ukladá do tabuľky, ktorá vo väčšine prípadov má 128 riadkov. V tejto tabuľke sa určuje poradie paketu v prvých 7 bitoch a posledný bit určuje, ktoré jadro procesora spracováva danú inštrukciu. Pokročilejšie sieťové karty umožňujú riadenie paketov podľa stanovených filtrov napríklad, posielanie špeciálnych paketov, ako je HTTP na porte 80 do vlastnej tabuľky na spracovanie, čím sa znižuje odozva systému a rýchlosť. Tieto filtre je možné následne konfigurovať pomocou ovládača v operačnom systéme s využitím funkcionality ethtool.

Pomocou technológie RSS poskytuje sieťový ovládač informácie operačnému systému o týchto HASH údajoch, ktoré sa následne spracovávajú na viacerých jadrách procesora. Samotná sieťová karta využívajúca túto HASH funkciu ktorá sa vypočítava pomocou rýchlych FPGA obvodov pomocou ktorých dokáže tento proces efektívne zrýchliť.

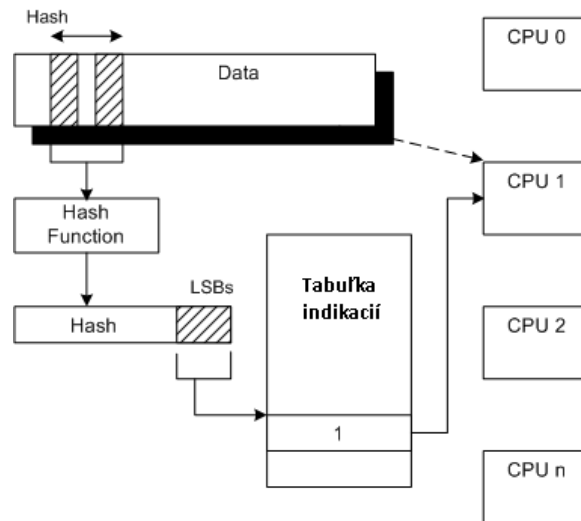


Fig. 4 Topológia RSS algoritmu s využitím HASH funkcií

Na Obr.4 je znázornený RSS algoritmus výberu jadra procesora. Pre výber daného jadra procesora sa využíva časť HASH hodnoty konkrétne LSB bit. Tento bit určuje smerovú hodnotu pridelenú vybranému jadrú procesora. Technológia RSS prináša tieto zlepšenia:

- Spracovanie prijatých paketov zo sieťového adaptéra na viacerých jadrách procesora.
- Dynamické vyváženie záťaže sieťovej karty a takisto aj procesora
- Pri komunikácií a rozložení na viaceré jadra procesora dochádza k zachovaniu poradia spracovávaných paketov.
- Opätovné načítanie Cache ktoré sa zdieľajú a vykonávajú na rovnakom jadre procesora.
- Algoritmus zvyšuje bezpečnosť systému pridaním HASHu zabezpečujúceho podpis daného paketu. Tento podpis chráni systém pred pokusmi zmanipulovať sieťovú prevádzku zasielaním podvrhnutých paketov.

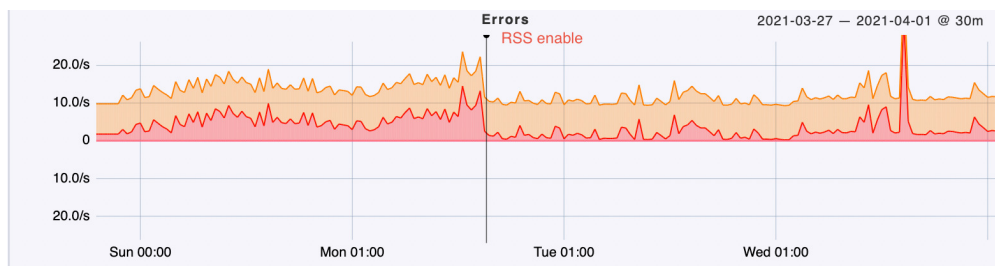


Fig. 5 Vplyv na stratu paketov pri použití RSS algoritmu

Na Obr.5 je znázornený graf reprezentujúci packet drop vznikajúci v kerneli operačného systému, ktorý nastával pred zapnutím technológie RSS a po zapnutí tejto technológie packet drop klesol. Z grafu vyplýva, že po zapnutí RSS technológie došlo k zníženiu vytvárania packet dropov na danom porte. Tieto packet dropy sa vytvárali hlavne s rastúcou sieťovou prevádzkou čiže technológia RSS umožnila monitorovať vyššiu sieťovú prevádzku pri rovnakom packet drope.

## V. ZHODNOTENIE

Začiatok práce bol venovaný nájdeniu vhodného mechanizmu ktorým by sa zastrešil zber údajov z dátovej siete. Boli otestované viaceré varianty, kde výsledkom bol protokol NetFlow s využitím python knižníc NFStream. Pomocou merania stability rýchlosti a modularity aplikácií vyšiel framework NFStram ako dostatočne modulárny aj pre budúce rozvíjanie, tak isto bolo možné pomocou tohto frameworku zakomponovať exportovanie dát do nerelačnej databázy Elasticsearch. Zápis do tejto databázy prebieha vytvoreným modulom, ktorý komunikuje pomocou JSON.

Pri použití navrhutej topológie merania dochádza k rozkladaniu výkonu sieťovej analýzy na viac jadier procesora. Pri neustále sa zvyšujúcej a premenlivej sieťovej prevádzke, ktorá je meraná, je potrebné rozkladať komunikáciu medzi sieťovou kartou a procesorom. Tento proces rozkladania rieši technológia Receive side scaling. Touto metódou sa docielilo rozloženie komunikácie so sieťovou kartou na viacero jadier. Rozloženie je možné využitím pomocou hardvérovej akcelerácie prípadne pomocou softvérovej akcelerácie. Využitím softvérovej akcelerácie nie je možné dosiahnuť vysoké zvýšenie výkonu sieťovej prevádzky. Najvhodnejšia metóda je využitie hardvérovej akcelerácie, kde časť prepočtu prevezme sieťová karta. Je možné využiť sieťovú kartu s technológiou FPGA, ktorá urýchľuje výpočty HASH funkcií. V niektorých sieťových kartách je možné sieťovú prevádzku triediť pomocou filtrov a posielat do jedného jadra iba pakety určitého typu. Týmto prístupom je možné znížiť frekvenciu obnovovania cache pamäte, keďže aplikácia využívajúca dáta beží na rovnakom jadre, kde dochádza ku komunikácií so sieťovou kartou.

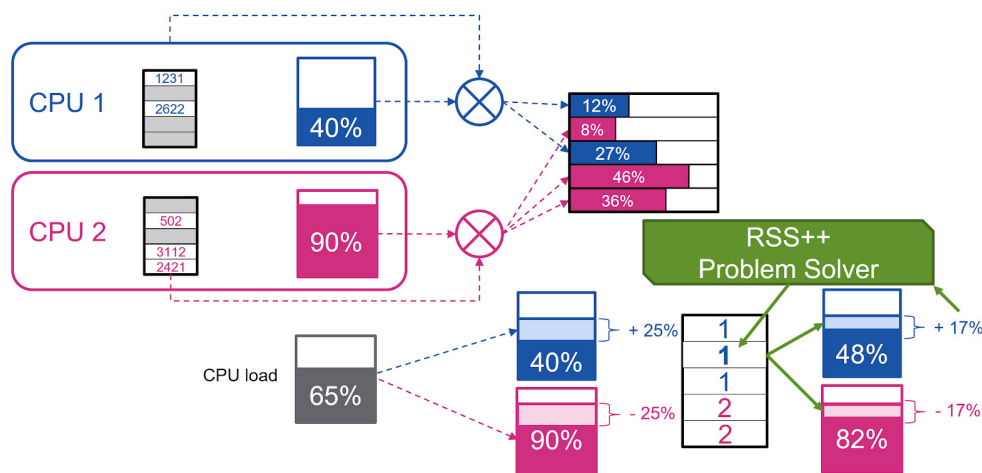


Fig. 6 Popis vylepšeného algoritmu RSS++

Použitím technológie RSS na rozloženie výkonu došlo k nerovnomernému zaťaženiu jednotlivých jadier procesora. Tento problém umožňuje preťažovanie jedného alebo viacerých jadier oproti iným čo vyplýva zo štruktúry samotného fungovania RSS. Na riešenie tohto problému sa časom našiel novší algoritmus RSS++, ktorý dokáže rovnomerne zaťažiť jadra procesora. Na Obr.6 je znázornená schéma algoritmu novej technológie RSS++, ktorá rieši problém preťažovania jednotlivých jadier[11]. Riešením tohto problému je pridelovanie paketov podľa vyťaženia daného jadra, ak je jedno jadro preťažované oproti ostatným, tak sa časť paketov, ktoré smerujú naň presunie na iné jadro, čím sa docieli vytvorenie rovnovážneho stavu. Tento rovnovážny stav umožňuje rovnomerne vyťažiť jadrá procesora, čím sa zvyšuje priepustnosť sieťovej karty, ale aj stabilita pri náhlom výkyve sieťovej prevádzky smerom nahor.

## VI. POĎAKOVANIE

Tento príspevok vznikol vďaka podpore Výskum v sieti SANET a možnosti jej ďalšieho využitia a rozvoja, kód výzvy OPVaI-VA/DP/2018/1.1.3-05, kód projektu: 313011W988.

## REFERENCIE

- [1] M. Stemm, R. Katz, and S. Seshan, "A network measurement architecture for adaptive applications," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064)*, vol. 1. IEEE, 2000, pp. 285–294.
- [2] B. Claise, G. Sadasivan, V. Valluri, and M. Djernaes, "Cisco systems netflow services export version 9," 2004.
- [3] U. D. Black and U. N. Black, *IP routing protocols: RIP, OSPF, BGP, PNNI, and Cisco routing protocols*. Prentice Hall Professional, 2000.
- [4] B. Trammell and E. Boschi, "An introduction to ip flow information export (ipfix)," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 89–95, 2011.
- [5] X. Yuan, C. Li, and X. Li, "Deepdefense: identifying ddos attack via deep learning," in *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 2017, pp. 1–8.
- [6] V. Duarte and N. Farruca, "Using libpcap for monitoring distributed applications," in *2010 International Conference on High Performance Computing & Simulation*. IEEE, 2010, pp. 92–97.
- [7] L. Deri, M. Martinelli, T. Bujlow, and A. Cardigliano, "ndpi: Open-source high-speed deep packet inspection," in *2014 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2014, pp. 617–622.
- [8] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.



- [9] A. Rucker, M. Shahbaz, T. Swamy, and K. Olukotun, "Elastic rss: Co-scheduling packets and cores using programmable nics," in *Proceedings of the 3rd Asia-Pacific Workshop on Networking 2019*, 2019, pp. 71–77.
- [10] S. Woo and K. Park, "Scalable tcp session monitoring with symmetric receive-side scaling," *KAIST, Daejeon, Korea, Tech. Rep*, 2012.
- [11] T. Barbette, G. P. Katsikas, G. Q. Maguire Jr, and D. Kostić, "Rss++ load and state-aware receive side scaling," in *Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies*, 2019, pp. 318–333.

# Rekurentné fuzzy neurónové siete pri riadení pohonov mobilných robotov

Ján VAŠČÁK

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

jan.vascak@tuke.sk

**Abstract** – Práca sa zaoberá využitím rekurentných neurónových sietí pri riadení kolesových pohonov mobilných robotov. Vychádzajúc z kinematického popisu robota je navrhnutý sériový regulátor ako kombinácia konvenčného fuzzy regulátora a neurónovej siete. Návrh bol simulačne testovaný a porovnávaný aj voči konvenčnej MLP sieti.

**Keywords** – fuzzy regulátor, mobilný robot, navigácia, rekurentná fuzzy neurónová sieť

## I. ÚVOD

Hlavnou témou tohto článku je navigácia dvojkolesového mobilného robota s využitím fuzzy rekurentných neurónových sietí (RFNN) vychádzajúcich z prác [1], [2]. Postupne sa venujeme analytickému popisu modelu mobilného robota v prostredí MATLAB a následne návrhu RFNN regulátora, ktorého cieľom je ovládať napätie jednosmerného motora kolesa robota.

Hlavnou motiváciou prepojenia fuzzy logiky a neurónovej siete je ich potenciál sa vzájomne dopĺňať svojimi vlastnosťami ako aj schopnosť obidvoch prostriedkov aproximovať analytické funkcie. Takýmto spôsobom máme možnosť zabezpečiť rýchlejšiu konvergenciu správania sa systému smerom k predpísanému stavu [3], [4].

## II. POPIS ŠTRUKTÚRY RFNN

Hlavnou nevýhodou klasických FNN je zameranie sa na konkrétne statické problémy kvôli svojej doprednej štruktúre. Preto pri našom type problému je vhodnejšie zaviesť RFNN založenú na supervíznom učení (učenie s učiteľom) a preto je vhodnejšia pre dynamické systémy. Tento typ sietí dokáže spracovať časovo premenlivý vstup a je aj možné dočasne ukladať informácie týmto spôsobom, čo vedie k vzniku jednoduchšej siete, ktorá ma menej prepojení a uzlov [5]. Ako je vidieť na Obr. 1, RFNN je tvorená klasickou FNN s tým, že v jej štruktúre je implementované spätnoväzobné spojenie v druhej vrstve. Táto úprava v konečnom dôsledku zabezpečuje pamäťové prvky a elimináciu časových problémov. Rekurentný neurón obsahuje vnútornú slučku znázornenú, ktorá zachytáva dynamickú odozvu systému [6].

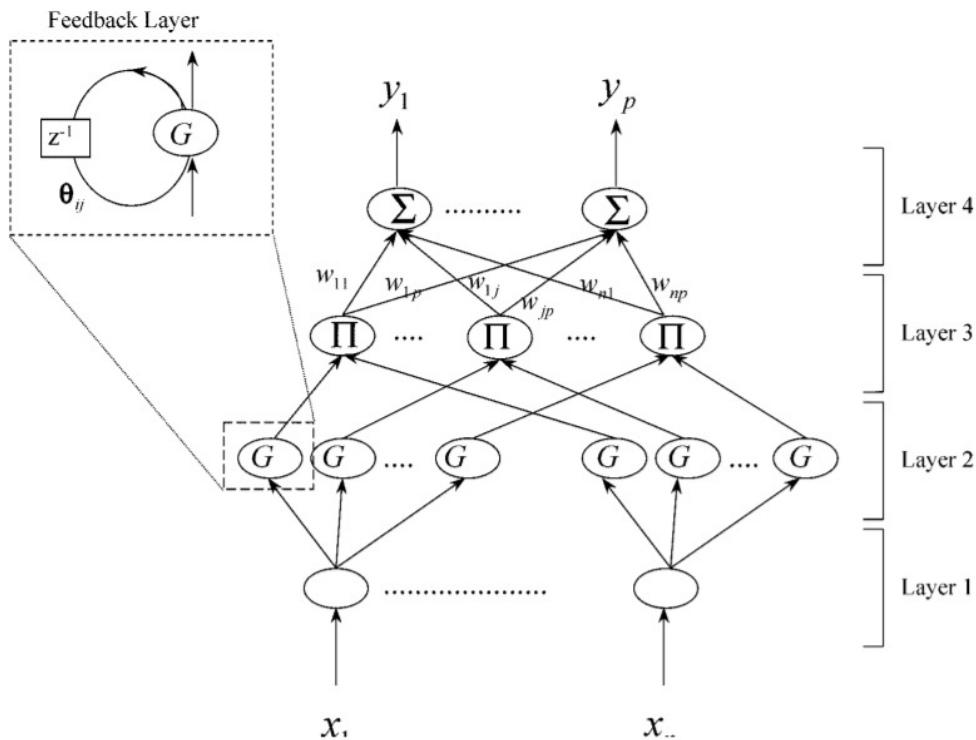
Štruktúra uvedená na Obr. 1 predstavuje fuzzy inferenčný systém, ktorý je využitý v našom riešení. Konkrétne, je táto tvorená  $n$  vstupov v prvej vrstve,  $m$  uzlov pre každú vstupnú vrstvu,  $p$  výstupných uzlov a  $m * n$  pravidiel uzlov. Samotný systém RFNN sa skladá zo 4 vrstiev a  $n + (n * m) + m + p$  uzlov, kde  $m$  je číslo pravidla. Ako bolo spomenuté vrstva 1 tvorí vstup RFNN, ktorá prijíma vstupné hodnoty. Vrstva 2 respektíve spätnoväzobná vrstva predstavuje výpočet gaussovských hodnôt. Vrstva 3 alebo fuzzy vrstva reprezentuje fuzzy pravidlá. Vrstva 4 je výstupom RFNN, kde každý uzol tvorí samostatný výstup systému. Prepojenia medzi 3. a 4. vrstvou sú tvorené váhovými hodnotami [1]. V ďalšom si uvedieme modifikáciu uvedeného princípu RFNN pre potreby nášho systému pre riadenie jednosmerných (DC) motorov robotického pohonu.

## III. POPIS RIADENIA DVOJKOLESOVÉHO ROBOTA

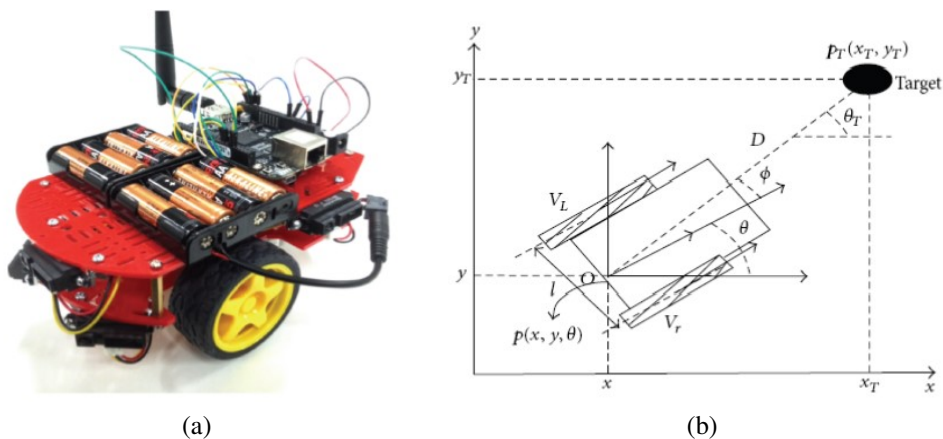
V našom projekte sme používali matematický model dvojkolesového mobilného robota QuickBot [4] s DC motormi, ktorý je znázornený na Obr. 2. Dynamický model bol odvodený analyticky a implementovali sme ho do programového prostredia MATLAB kde sme vykonali všetky simulácie navigačného riadenia.

Lineárny dynamický systém mobilného robota s DC motormi má v maticovom zápise nasledujúci tvar:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \quad (1)$$



Obr. 1 Topologická štruktúra RFNN [1].



Obr. 2 Mobilný robot QuickBot a schéma jeho dynamického modelu.

kde matica dynamiky  $A$  mobilného robota s DC motormi je definovaná nasledovne:

$$\begin{bmatrix} -2\frac{Bze}{Mr^2} & 0 & kt\frac{N}{Mr} & kt\frac{N}{Mr} \\ 0 & -2Bze\frac{b^2}{Ir^2} & ktN\frac{b}{Ir} & ktN\frac{b}{Ir} \\ -\frac{60}{2\pi}ke\frac{N}{rLr} & -\frac{60}{2\pi}keN\frac{b}{rLr} & -\frac{Rv}{Lr} & 0 \\ -\frac{60}{2\pi}ke\frac{N}{rLr} & \frac{60}{2\pi}keN\frac{b}{rLr} & 0 & -\frac{Rv}{Lr} \end{bmatrix}. \quad (2)$$

Rovnica (3) predstavuje stavový vektor, ktorý obsahuje štyri stavové premenné, z toho prvé dve sú lineárna a uhlová rýchlosť robota a ďalšie dve sú prúd pretekajúci pravým a ľavým DC motorom (vid' Obr. 2(b)).

$$\mathbf{x}^T = [v \quad \omega \quad i_r \quad i_l]. \quad (3)$$

V rovnici (4) je definovaný vektor vstupu, ktorý obsahuje napätia na svorkách pravého a ľavého motora.

$$\mathbf{u}^T = [u_R \quad u_L]. \quad (4)$$

Z dynamického modelu je možné získať aktuálnu lineárnu a uhlovú rýchlosť. Na to, aby sme poznali aktuálnu pozíciu robota, je potrebný kinematický model, ktorý je popísaný vzťahom (5).

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi & 0 \\ \sin \varphi & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}. \quad (5)$$

Polohové súradnice potom získame ako riešenie sústavy diferenciálnych rovníc prvého rádu:

$$\mathbf{q}^T = [x \quad y \quad \varphi]. \quad (6)$$

Na odvodenie pohybových rovníc je možné použiť Lagrangeove rovnice II. rádu, kde je potrebné si najprv určiť vektor  $n$  nezávislých zovšeobecnených súradníc  $\mathbf{q}$  a zostaviť Lagrangeovu funkciu  $L(t)$ , ktorú získame ako rozdiel celkovej kinetickej a potenciálnej energie systému. Výsledkom toho je nasledujúcich  $n$  pohybových rovníc:

$$\frac{d}{dt} \left( \frac{\partial L(t)}{\partial \dot{q}_i(t)} \right) - \frac{\partial L(t)}{\partial q_i(t)} + \frac{\partial D(t)}{\partial \dot{q}_i(t)} = Q_i, \quad (7)$$

kde  $Q_i$  je vstup pôsobiaci na  $i$ -tú zovšeobecnú súradnicu. Bližší postup odvodzovania systému (7) je možné nájsť napr. v [7].

Model robota bude napokon riadený prostredníctvom regulátora, ktorého vstupmi sú referenčná pozícia, ktorú nastavuje používateľ, aktuálna rýchlosť a aktuálna pozícia, ktoré sú získané priamo z mobilného robota. Na základe týchto informácií musí regulátor vygenerovať napätie na svorkách motorov mobilného robota, ktoré vyvolá pohyb robota smerom k požadovanému cieľu.

#### IV. NÁVRH RIEŠENIA

Na splnenie nášho cieľa sme použili dva regulátory zapojené v sérii. Prvý je regulátor kinematiky, ktorého úloha je generovanie referenčnej rýchlosti na základe pozičných odchýlok, ktoré sú vypočítané ako rozdiel aktuálnej a referenčnej pozície. V prípade, že je v trase prekážka, systém sa prepne do režimu vyhýbania a referenčná pozícia v ďalšom kroku sa zmení na takú hodnotu, aby robot bezpečne obchádzal prekážku. V prípade, že prekážka bola bezpečne obídená, referenčná pozícia sa opäť prepíše na pôvodnú hodnotu zadanú používateľom. Tento regulátor bol implementovaný v podobe Mamdaniho fuzzy regulátora.

Druhý regulátor je regulátor dynamiky. Jeho úlohou je na základe aktuálnej a požadovanej rýchlosti generovať napätie na svorkách DC motora. Tento regulátor bol implementovaný v podobe neurónovej a fuzzy neurónovej siete, ktoré boli natrénované ako inverzné modely pre potreby ďalšieho porovnania. Inými slovami, tento typ sériového regulátora ako celok zahŕňa dva stupne hybridizácie. Prvý stupeň je rozdelenie regulátora na dve časti, z ktorých je každá realizovaná inou metódou výpočtovej inteligencie. Druhý stupeň predstavuje realizácia regulátora dynamiky pomocou hybridnej fuzzy neurónovej siete.

##### A. Regulátor kinematiky

Úlohou tohto regulátora je generovanie referenčnej rýchlosti (vstup do dynamického modelu) na základe pozičnej odchýlky tak, aby sme tieto odchýlky vedeli minimalizovať. V polárnom súradnicovom systéme je poloha referenčného bodu vyjadrená pomocou euklidovskej vzdialenosti  $\rho$  a odchýlky smerového natočenia  $\alpha$ :

$$\rho = \sqrt{\Delta x^2 + \Delta y^2}, \quad (8)$$

$$\alpha = \arctan \frac{\Delta y}{\Delta x}. \quad (9)$$

Pre jeho implementáciu bol použitý Mamdaniho regulátor s defuzzifikáciou podľa ťažiska. Regulátor má dva vstupy a dva výstupy. Vstupy predstavujú pozičné odchýlky  $\alpha$  a  $\rho$  a výstupy sú referenčné uhlové rýchlosti pravého a ľavého kolesa  $\omega_R$  a  $\omega_L$ . Všetky fuzzy množiny boli rovnomerne rozdelené tak, aby pokrývali celé univerzum vstupných a výstupných veličín. Fuzzy lingvistická premenná  $\rho$  je definovaná na univerze euklidovskej vzdialenosti robota a referenčnej pozície. Táto lingvistická premenná je rozdelená medzi tromi slovnými hodnotami *Low*, *Medium* a *High*. Lingvistická premenná  $\alpha$  je definovaná na univerze uhlovej odchýlky smerovania robota od cieľa a nadobúda taktiež tri slovné hodnoty *Negative*, *Neutral* a *Positive*. Výstupné premenné regulátora kinematiky  $\omega_R$  a  $\omega_L$  sú definované na univerze uhlových rýchlostí ľavého a pravého kolesa a môžu nadobúdať päť slovných hodnôt *Very Low*, *Low*, *Medium*, *High* a *Very High*.

Báza pravidiel, ktorá vyjadruje vzťahy medzi týmito premennými za účelom aproximácie vzťahov medzi vstupmi a výstupmi regulátora bola navrhnutá nasledovným spôsobom:

- 1) AK  $\rho = H$  a  $\alpha = +$  Potom  $\omega_R = VH$  a  $\omega_L = M$
- 2) AK  $\rho = H$  a  $\alpha = N$  Potom  $\omega_R = H$  a  $\omega_L = H$

- 3) AK  $\rho = H$  a  $\alpha = -$  Potom  $\omega_R = M$  a  $\omega_L = VH$
- 4) AK  $\rho = M$  a  $\alpha = +$  Potom  $\omega_R = H$  a  $\omega_L = L$
- 5) AK  $\rho = M$  a  $\alpha = N$  Potom  $\omega_R = M$  a  $\omega_L = M$
- 6) AK  $\rho = M$  a  $\alpha = -$  Potom  $\omega_R = L$  a  $\omega_L = H$
- 7) AK  $\rho = L$  a  $\alpha = +$  Potom  $\omega_R = M$  a  $\omega_L = VL$
- 8) AK  $\rho = L$  a  $\alpha = N$  Potom  $\omega_R = L$  a  $\omega_L = L$
- 9) AK  $\rho = L$  a  $\alpha = -$  Potom  $\omega_R = VL$  a  $\omega_L = M$

### B. Regulátor dynamiky

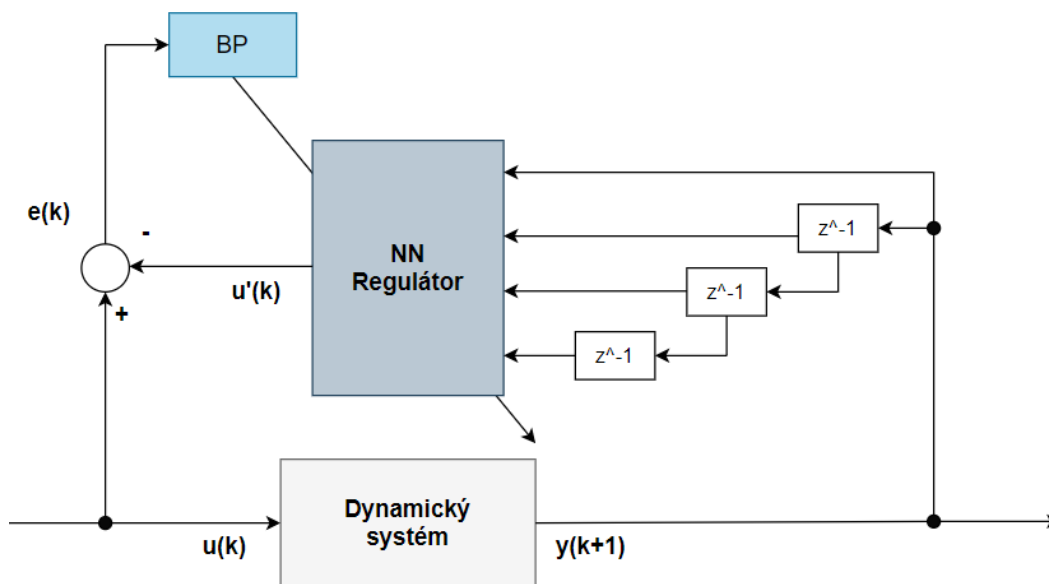
Regulátor dynamiky má za úlohu sledovanie referenčnej rýchlosti. To znamená, že musí generovať taký akčný zásah (v našom prípade napätia motorov), aby výstup dynamického systému sledoval referenčné vstupy do regulátora dynamiky, ktoré boli vygenerované kinematickým regulátorom. Pre tento účel bola použitá RFNN, ktorej štruktúra bola popísaná v štádi II a taktiež pre potreby porovnávania klasický MLP s jednou skrytou vrstvou.

Na rozdiel od [1], kde túto sieť trénovali ako dopredný model, bola v tomto prípade sieť trébovaná ako inverzný model, ktorý bol následne použitý ako dopredný regulátor. Tým istým spôsobom bola trébovaná aj sieť MLP. Tak ako všetky neurónové siete, aj táto pracuje v režime trébovania a režime testovania.

V režime trébovania sa sieť snaží aproximovať inverzný dynamický model mobilného robota, čiže výstupy zo systému predstavujú vstupy do siete a výstupy zo siete zasa vstupy do systému. V tomto prípade sa však na vstup siete okrem aktuálneho výstupu systému privádzajú aj výstupy niekoľko krokov dozadu.

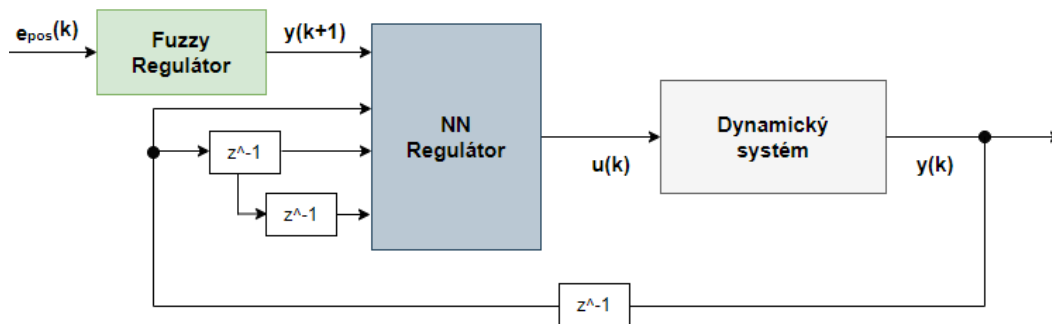
Obr. 3 predstavuje schému zapojenia RFNN/NN siete vo fáze učenia. Celý proces učenia je popísaný v nasledujúcich krokoch:

- 1) Na systém je privedený vstup  $u(k)$ , ktorý v ňom vyvolá odozvu  $y(k+1)$ .
- 2) Na vstup siete je privedený výstup systému  $y(k+1)$  spolu s výstupmi 1–3 kroky do zadu.
- 3) Sieť na základe vstupov vygeneruje výstup  $u'(k)$ , ktorý je aproximáciou reálneho vstupu do systému.
- 4) Na základe rozdielu reálnej a aproximovanej hodnoty vstupu sa s využitím algoritmu spätného šírenia chyby adaptujú parametre siete.



Obr. 3 RFNN vo fáze učenia.

V režime testovania plní natrébovaný inverzný model funkciu regulátora. Blokovaná schéma jeho zapojenia je zobrazená na obrázku 4. Rovnako ako v prípade trébovania sú na vstup regulátora privedené oneskorené hodnoty výstupov z dynamického systému. Rozdiel je iba v hodnote  $y(k+1)$ , ktorá nemôže byť v tomto prípade privedená z dynamického systému, nakoľko je táto hodnota reakciou na vstup  $u(k)$ . Preto je táto hodnota získaná z kinematického fuzzy regulátora a predstavuje referenčnú hodnotu, na ktorú by sa v ďalšom kroku mala výstupná hodnota dostať po aplikovaní vstupu z inverzného modelu.



Obr. 4 RFNN vo fáze testovania.

## V. EXPERIMENTY A ICH VYHODNOTENIE

Experimentálna časť je zameraná na určenie presnosti trénovaných modelov, s ktorou dokázali aproximovať inverznú dynamiku mobilného robota. Tabuľka 1 znázorňuje porovnanie strednej kvadratickej odchýlky na testovacej množine, ktoré dosiahli MLP a RFNN.

Tabuľka 1  
 Porovnanie strednej kvadratickej odchýlky MLP a RFNN pre rôzne kombinácie ich parametrov.

Počet neurónov	Parameter učenia	MLP	RFNN
10	0.001	$3.257 \times 10^{-5}$	$3.02 \times 10^{-5}$
10	0.002	$2.50 \times 10^{-5}$	$2.19 \times 10^{-5}$
10	0.005	$5.02 \times 10^{-5}$	$4.02 \times 10^{-5}$
10	0.01	$9.20 \times 10^{-5}$	$8.20 \times 10^{-5}$
10	0.05	$5.20 \times 10^{-4}$	$4.33 \times 10^{-4}$
11	0.001	$5.60 \times 10^{-5}$	$5.50 \times 10^{-5}$
11	0.002	$6.66 \times 10^{-4}$	$5.56 \times 10^{-4}$
11	0.005	$6.55 \times 10^{-5}$	$6.25 \times 10^{-5}$
11	0.01	$7.55 \times 10^{-5}$	$7.00 \times 10^{-5}$
11	0.02	$6.34 \times 10^{-5}$	$6.22 \times 10^{-5}$
12	0.001	$5.56 \times 10^{-4}$	$5.45 \times 10^{-4}$
12	0.002	$7.23 \times 10^{-5}$	$6.43 \times 10^{-5}$
12	0.005	$3.31 \times 10^{-4}$	$2.21 \times 10^{-4}$
12	0.01	$5.35 \times 10^{-5}$	$4.33 \times 10^{-5}$
12	0.02	$3.57 \times 10^{-5}$	$4.22 \times 10^{-5}$
13	0.001	$1.36 \times 10^{-5}$	$1.22 \times 10^{-5}$
13	0.002	$4.62 \times 10^{-5}$	$2.22 \times 10^{-5}$
13	0.005	$1.35 \times 10^{-5}$	$1.10 \times 10^{-5}$
13	0.01	$6.67 \times 10^{-5}$	$5.87 \times 10^{-5}$
13	0.02	$1.38 \times 10^{-5}$	$1.23 \times 10^{-5}$

V prípade MLP sa najlepšie výsledky dosiahli pri 13 neurónoch a parametre učenia 0,005. Podobne, pri rovnakých parametroch aj RFNN dosiahla najlepšie výsledky. Pri vzájomnom porovnaní vychádza RFNN o niečo lepšie ako MLP.

## VI. ZÁVER

Ako experimenty potvrdzujú, natrénovaný inverzný model spoľahlivo plní funkciu regulátora. Všetky simulácie prebehli bezproblémovo a vďaka hybridnému modelu fuzzy rekurentnej neurónovej siete bolo možné v 2D priestore navigovať dvojkolesový robot. Z výsledkov je jasné, že lepšie hodnoty boli zaznamenané pri RFNN, čím sa dá táto metóda považovať za vhodnú pre dané riešenie.

## LITERATÚRA

- [1] C.-H. Lee and C.-C. Teng, "Identification and control of dynamic systems using recurrent fuzzy neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 4, pp. 349–366, 2000.
- [2] J. Vaščák, P. Papcun, I. Zolotová, and M. Puheim, "Intelligent space design for robotic and IoT applications," in *2020 Cybernetics Informatics (KI)*, 2020, pp. 1–6.
- [3] K. ichi Funahashi and Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural Networks*, vol. 6, no. 6, pp. 801–806, 1993.
- [4] H. Omrane, M. S. Masmoudi, and M. Masmoudi, "Fuzzy logic based control for autonomous mobile robot navigation," *Comput. Intell. Neurosci.*, vol. 2016, pp. 9 548 482:1–9 548 482:10, 2016.

- [5] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Computation*, vol. 1, no. 2, pp. 270–280, 1989.
- [6] B. Šter, "Recurrent neural networks in a mobile robot navigation task;" in *Artificial Neural Nets and Genetic Algorithms*, V. Kůrková, R. Neruda, M. Kárný, and N. C. Steele, Eds. Vienna: Springer Vienna, 2001, pp. 169–172.
- [7] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to Autonomous Mobile Robots, Second Edition*, ser. Intelligent robotics and autonomous agents. MIT Press, 2011.

# Viacpreskoková mobilné siete MANET a mobilné siete 6. generácie

<sup>1</sup>Gabriela VAŠKOVÁ, <sup>2</sup>Ján PAPAJ

<sup>1 2</sup> Katedra elektroniky a multimediálnych technológií,  
Fakulta elektrotechniky a informatiky Technická Univerzita v Košiciach,  
Slovenská republika

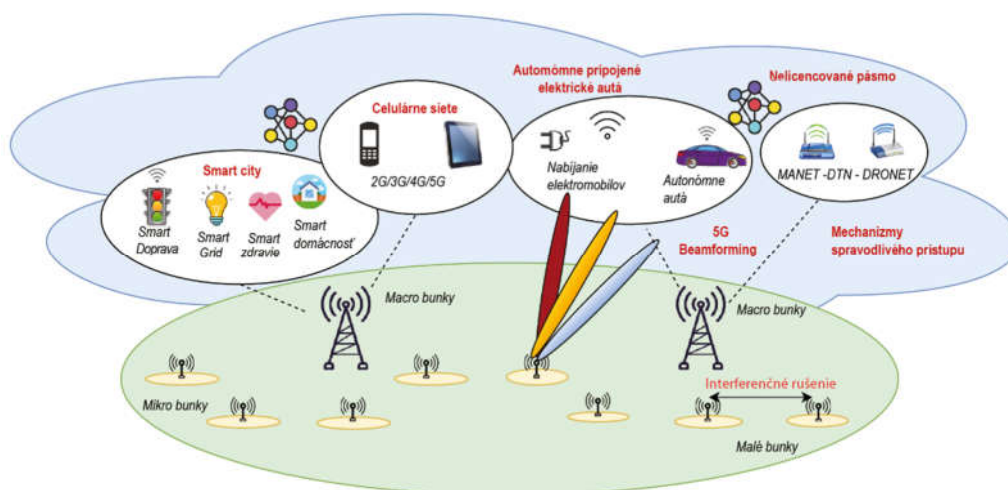
<sup>1</sup>gabriela.vaskova@student.tuke.sk, <sup>2</sup>jan.papaj@tuke.sk

**Abstrakt** — V súčasnej dobe sa do popredia dostáva praktická otázka poskytovania časovo citlivých služieb prostredníctvom nových bezdrôtových technológií a nových sietí šiestej generácie. Tieto siete budú využívať rozličné typy viacpreskokových ad-hoc sietí (Mobile Ad-Hoc Network MANET, Drones network DRONET, Delay Tolerant Network DTN). Kľúčovým elementom týchto sietí bude teda smerovací protokol Jeho hlavná úloha je poskytnutie QoS služieb a zabezpečenie sieťovej výkonnosti. Výber optimálneho smerovacieho protokolu a vhodných parametrov pre danú sieť nakoľko tieto parameter dokážu ovplyvniť správanie sa celej siete. Tento článok sa zaoberá základnou možnosťou využívania MANET sietí v mobilných sieťach 6. generácie.

**Kľúčové slová** — 6G siete, Smerovanie, MANET, strojové učenie, Blockchain

## I. ÚVOD

Šiesta generácia (6G) je označenie pre šiestu generáciu bezdrôtových a mobilných technológií, ktorá nadväzuje na 4G a 5G siete. Výskum oblastí sietí 6G je stále ešte v počiatočnej fáze, a preto sa nedá jednoznačne povedať, ako budú tieto siete vyzerajú. Zatiaľ to predstavuje iba súbor teórií o tom, ako by mohli mobilné zariadenia, automobily a ďalšie veci, ktoré budú spojené cez Wi-Fi vyzerajú. Je to budúcnosť bezdrôtového internetu (Obr. 1). Vo výskumných prácach sú 6G siete popisované ako plne integrovaný internetový systém, ktorý umožňuje okamžitú komunikáciu medzi používateľmi, zariadeniami a okolitým prostredím [1].



Obr. 1 Základná idea sietí novej generácie - 6G.

V súčasnosti IoT (Internet of Things) nám umožňuje ovládať zariadenia v inteligentnej domácnosti s využitím smartfónov a nových mobilných technológií. Vývoj technológií smeruje k tomu, aby sme dosiahli IoE (Internet of Everything). To však závisí aj od vývoja nových mobilných sietí šiestej generácie [4]. Ak odborníci hovoria o 6G sieťach, tak často hovoria o [5]:

- **Rýchlosť 1 TB/s** – 6G siete umožnia dosiahnuť rýchlosť na internetovom zariadení až 1 Terabajt za sekundu. To je tisíckrát rýchlejšia ako 1 Gb/s, čo je zatiaľ najvyššia rýchlosť,



aká je dostupná vo väčšine internetových domácnosti.

- *Umelá inteligencia* – Už pri 5G sieťach sa hovorilo o automatizovaných automobiloch a dronoch, ktoré by pomocou diaľkového ovládania riadili továrne a taktiež aj o inom využití Umelej inteligencie (Artificial Intelligence AI). Odborníci si myslia, že na zabezpečenie bezproblémového chodu 6G technológií bude potrebná Umelá inteligencia.

System 6G zahŕňa širokú škálu rôznych zdrojov odlišných inteligentných aplikácií, ktoré generujú obrovské množstvo dát. Vysokovýkonné výpočty a kvantové výpočty musia byť použité na inteligentnú analýzu dát. Očakáva sa, že výpočtová technika spôsobí revolúciu v oblasti výpočtovej techniky povolením vyšších rýchlostí, aké používatelia doteraz ešte nezažili. Kľúčová vlastnosť kvantovej komunikácie sú zabezpečené kanály, kde každý kanál nesie odlišné bezpečnostné protokoly zabudované do šifrovaných údajov. Tieto vlastnosti zabezpečenia navyše k vysokej rýchlosti robia kvantové výpočty výhodné pre bezpečné 6G inteligentné aplikácie [4].

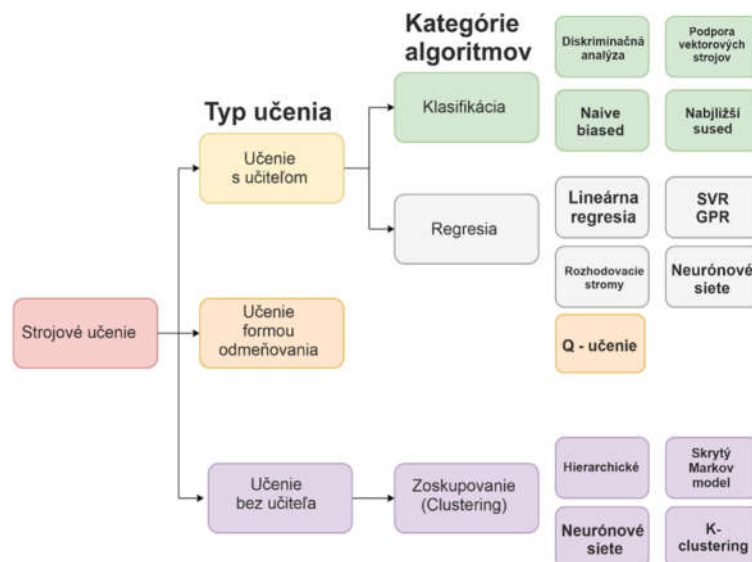
Medzi komponenty, ktoré 6G siete patria sú *edge computing*, *cloud computing*, *AI*, *blockchain*, *network slicing*, a podobne. V sieťach piatej generácie bolo navrhnuté segmentovanie siete, ale realizácia sa očakáva až v sieťach šiestej generácie. Network slicing založené na Software-defined networking (SDN) a Network function virtualization (NFV), zdieľa fyzické zdroje na povolenie rôznych častí aplikácie. Proces network slicing zahŕňa optimalizáciu rôznych sieťových parametrov [4].

Mobilné ad-hoc siete (MANET) sú bezdrôtové, autonómne samo-organizujúce sa siete bez pevnej topológie. V takejto sieti, slúži uzol v rovnaký čas aj ako vysielač aj ako prijímač. Všetky uzly sú si rovnocenné a môžu sa voľne pohybovať, pripojiť či odpojiť zo siete. Mobilné uzly, ktoré sú v rádiovom dosahu s inými uzlami môžu priamo komunikovať a posilať nevyhnutné informácie. Na takúto komunikáciu majú všetky uzly vstavané bezdrôtové rozhranie. Takýto druh siete je plne distribuovaný a môže fungovať na akomkoľvek mieste bez pomoci akejkoľvek pevnej infraštruktúry akou sú napríklad prístupové body alebo základňové stanice. Rozdiel senzorovými (WSN) a MANET sieťami je v tom, že pri MANET sieťach je zmena topológie väčšinou zapríčinená zmenou polohy avšak pri WSN je zmena topológie väčšinou zapríčinená zlyhaním senzora respektíve vybitím batérie a taktiež v sieti neexistuje žiadny zberný uzol, ktorý je väčšinou statický a len funguje ako výstupná brána pre zozbierané údaje. Takéto MANET siete majú najlepšie využitie v oblastiach, kde je zložitá geografická prostredie, ktoré vyžaduje mobilné uzly bez pevnej infraštruktúry [2]. Príkladom môžu byť núdzové situácie, kde záchranné služby nemôžu využiť pôvodnú rádiovú sieť, či už z dôvodu preťaženia siete, jej zničenia, alebo slabého signálu v oblasti. Avšak dynamickosť MANET sietí, frekventované zmeny v sieti zapríčinené mobilitou, otvorené bezdrôtové vysielanie robia tieto siete náchylné voči bezpečnostným útokom. Tento problém vychádzajúci zo základných charakteristík MANET sietí a je neustále testovaný novými bezpečnostnými protokolmi, ktoré sa snažia zachovať výhody týchto sietí, ale nie na úkor ich bezpečnosti [3].

## II. PROCES SMEROVANIA V 6G SIETI

Nepretržitý nárast dátovej prevádzky znamená potrebu inteligentných technológií na sieťovej vrstve. 6G siete môžu pomôcť dosiahnuť autokonfiguráciu, samoorganizáciu, „samoliečenie“, samooptimalizáciu a zotavenie po poruche. Inteligentne plánujú, optimalizujú a vyberajú najvhodnejšia akcia po využití vhodných poznatkov [4]. Teda využitie 6G bude viesť k zvýšeniu uskutočniteľnosti, skráteniu času na zotavenie a zlepšeniu poskytovania kvality služieb pre konečného spotrebiteľa [5].

S hlbšou konvergenciou mobilných sietí by mala slúžiť 6G pre širokú škálu aplikácií v rôznych scenároch. Existuje obrovský rozdiel v požiadavkách pre rôzne služby. Očakáva sa, že 6G spôsobí revolúciu od „prepojených vecí“ po „prepojenú inteligenciu“, ktoré môžu autonómne a inteligentne vyvinúť sieť podľa rôznych požiadaviek na službu [4].



Obr. 2 Strojové učenie preMANET v 6G sieti.

Umelá inteligencia (AI) je odvetvie počítačovej vedy, ktoré umožňuje strojom myslieť a konať ako ľudia. Zahŕňa technológie ako strojové učenie, hlboké učenie, teória optimalizácie, teória hier a evolučné algoritmy. Strojové učenie je založené na princípe, že stroj sa môže učiť identifikáciou vzorcov z daných údajov a následne sa podľa toho rozhodnúť s minimálnym ľudským zásahom. Pri riešení súčasných problémov sa používajú techniky strojového učenia ako sú kontrolované učenie, nekontrolované učenie a posilňovacie učenie, ktoré zvyšuje jeho efektívnosť v komunikačných systémoch prostredníctvom 6G sietí [6].

6G bude čoraz viac flexibilné so sieťovým vybavením na podporu rôznych aplikácií. Sieťová architektúra by mala byť otvorená, aby pri virtualizácii siete požadovaná obslužná sieť bola schopná vygenerovať ľubovoľnú aplikáciu zdieľanej fyzickej infraštruktúry. Medzitým rozhranie by malo byť otvorené na podporu interoperability rôznych účastníkov a flexibilný harmonogram virtualizovaných sieťových prvkov [7].

#### A. Blockchain v MANET a 6G

V nových sieťach bude podporovaná široká škála služieb a ich požiadavky sa budú dynamicky meniť. S konvergenciou 3C zdroje možno použiť na podporu pokročilých výpočtových technológií technológie ako blockchain. Blockchain je kombinácia viacerých výpočtových technológií vrátane distribuovanej databázy (t. j. ukladanie do pamäte cache), inteligentnej zmluvy a mechanizmu konsenzu (t. j. výpočty). Blockchain je decentralizovaná účtovná kniha zabezpečená proti neoprávnenej manipulácii, ktorú je možné použiť pre zaznamenávanie a aktualizáciu využitia dynamického spektra. Zdieľanie distribuovaného spektra teda umožňuje server blockchain, buď registrácia využívania spektra alebo uvoľnenie. Dá sa vykonať celkom rýchlo, bez toho, aby bolo potrebné sa spoliehať na jednotku centralizovanej správy spektra [].

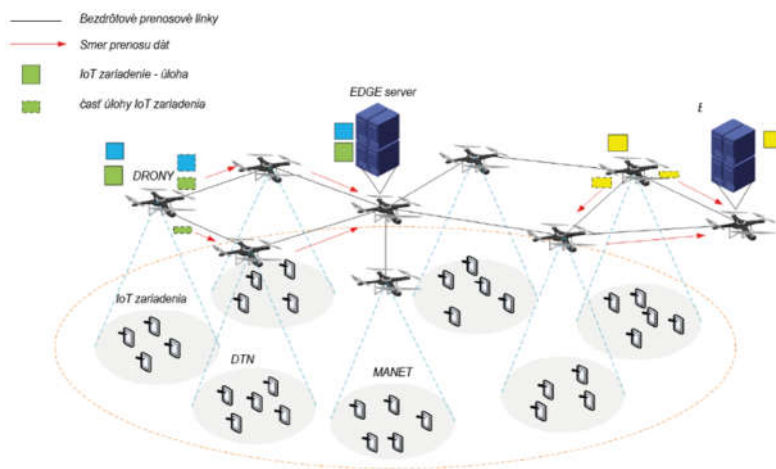
6G sieť sa snaží rozvíjať vysoko dynamicky a inteligentný systém, ktorý umožňuje sieti zmeniť prostredie tak, aby vyhovovalo rôznym aplikačným požiadavkám a typom služieb, napríklad vylepšené mobilné širokopásmové pripojenie (eMBB), vysoko spoľahlivé, s nízkou latenciou komunikácie (uRLLC) a masívny typ strojovej komunikácie (mMTC) [4].

Použitie malých mobilných uzlov, ako sú napríklad drony, na zber dát má niekoľko výhod. Medzi základné výhody patrí širšie pokrytie oblasti ako jedným väčším uzlom, sú výrazne lacnejšie, majú ľahšie nasadenie a poskytujú redundanciu. Úspešná smerovacia schéma očakáva, že zlyhanie týchto malých uzlov je možné, vzhľadom na relatívne nízku cenu uzlov a môže mať úlohy, pri ktorých uzol môže byť poškodený alebo stratený. Okrem toho nové uzly môžu byť nasadené kedykoľvek [2].

Blockchain je zdieľaná distribuovaná databáza, v ktorej dáta, respektíve záznamy sú uložené v blokoch. Blockchain je nový režim integrovanej aplikácie distribuovaných dátových úložísk, prenos z bodu do bodu, mechanizmus konsenzu, šifrovací algoritmus a ďalšie počítačové technológie. V zásade je blockchain založený na tom, že môže vyriešiť problém informačnej asymetrie a dosiahnuť kooperatívnu dôveru a spoločné konanie medzi viacerými subjektmi. Okrem toho môže slúžiť na vytvorenie bezpečnostných systémov, ako je autentifikácia identity systému na zvýšenie bezpečnosti bezdrôtovej 6G siete [4], [5], [6].

### B. Smerovanie a strojové učenie v MANET

Proces smerovania vo všeobecnosti venuje problematike doručovaniu paketov zo zdroja do cieľového miesta s optimálnym využitím sieťových zdrojov, pri poskytovaní QoS (Quality of Service a Quality of Experiences) (Obr. 3).



Obr. 3 MANET a jeho aplikácia v 6G sieti.

Proces optimalizácie komunikačnej cesty sa tradične zaoberá aj problémom nestability komunikačnej cesty prostredníctvom:

- Optimalizácie komunikačných ciest vzhľadom na predchádzajúce komunikačné podmienky s cieľom efektívne predikovať komunikačné cesty* – smerovací protokol sa v MANET snaží o optimalizovanie smerovania na základe aktuálnych prenosových podmienkach.
- Optimalizáciu komunikačných ciest s dôrazom na rozsahom komunikačných scenárov* – optimalizácia MANET smerovacích protokolov pre široké spektrum prenosových podmienok za účelom dosiahnutia požadovaných prenosových vlastností.

Smerovacie protokoly sa v posledných rokoch vyvinuli a prispôbili sa moderným MANET sieťam vo sfére IoT. Strojové učenie (Machine learning) vo všeobecnosti využíva také algoritmy, ktoré sa dokážu s plynutím času postupným zberom dát vylepšovať. Práve táto vlastnosť sebavylepšovania je využiteľná v procese hľadania optimálnej komunikačnej cesty medzi zdrojovým a cieľovým mobilným uzlom. Efektívnosť smerovacieho protokolu primárne závisí od dátového kontextu, ako aj od aplikácie a môže byť rozdelené viacerými spôsobmi. Niektoré z týchto delení zahŕňajú smerovanie založené na geografickej polohe, smerovanie založené na QoS a medzivrstvové (cross-layer) smerovanie. Strojové učenie si elegantne našlo svoju cestu v tejto doméne poskytovaným mocným nástrojom na riešenie niektorých problémov spojených s návrhom smerovacích algoritmov [2].

Tabuľka 1

Algoritmy využívajúce učenie s učiteľom pre MANET siete[4]

Technika	Charakteristiky / Výhody	Limitácie / Nevýhody
Umelá neurónová sieť	- Ťažkosti s aktualizáciou nových informácií - Neschopnosť odhaliť neobvyklé útoky - Založené na princípe ľudského mozgu - Schopný pracovať v limitovanom, hlučnom a nekompletnom prostredí	- Vysoké procesné nároky - Pomalý učiaci proces, nevhodný pre dynamické siete
Rozhodovací strom	- Vhodné pre rozmerne množiny údajov - Vysoká presnosť detekcie	- Vyžaduje vysoký výpočtový výkon
Bayesová sieť	- Rýchly, škálovateľný, používa neparametrický získavanie údajov - Vie pracovať s nekompletnými množinami dát	- Vyžaduje vysoký výpočtový výkon - Nemusi správne fungovať ak je zle natanené počiatkové vedomosti
Podpora vektorového stroja	- Schopný riešiť nelineárne klasifikácie a regresie - Automaticky sa vyhýba preplňovaniu údajov	- Veľký počet zle klasifikovaných dát - Trénovanie zaberá väčší čas

Tabuľka 2

Algoritmy využívajúce učenie bez učiteľa pre MANET siete [2]

Technika	Charakteristiky / Výhody	Limitácie / Nevýhody
K-Clustering	-Vysoká pravdepodobnosť úspešného pridelenia do správnej clustera -Môže byť použité na veľkú množinu dát	-Prítomnosť odľahlých hodnôt môže viesť k nesprávnemu klasifikovaniu
Učenie pomocou asociačného pravidla	-Používaný na extrahovanie dôležitých pravidiel, ktoré sa snažia nájsť spoločné charakteristiky v množine dát	-Časovo neefektívny
Skrytý Markov model	-Použitá Stochastická Markova teória -Topológia a schopnosti systému sú prepojené cez tranzitné pravdepodobnosti	-Výpočtovo náročný
Genetický algoritmus	-Založený na biologickej evolučnej teórii -Evolúcia pomocou selekcie, rekombinácie a mutácie -Vhodné pre veľké dátové množiny	-Vysoká pravdepodobnosť mutácie -Nie je zaručená časová odozva
Učenie nejasnou logikou	-Založený na odhade a neistote	-Náročné aktualizovanie identifikácií a pravidiel počas spustenia

Algoritmy môžu riešiť tri základné úlohy a to klasifikáciu, regresiu alebo zhlukovanie (Tab.1 a Tab.2). Na základe vykonávaných úloh sa tieto algoritmy delia do troch kategórií. Ide o učenie s učiteľom, učenie bez učiteľa a učenie formou odmeňovania s ktorým budeme pracovať aj my v tejto práci. Dynamický charakter MANET sietí si vyžaduje také smerovacie algoritmy, ktoré budú udržiavať minimálne oneskorenie doručenia dátových paketov aj v prípadoch nízkeho a vysokého zaťaženia siete (Obr.3). Strojové učenie má v MANET optimalizovať nielen samotný prenos dát, ale aj proces výberu optimálnej komunikačnej cesty v zmysle minimalizovania end-to-end oneskorenia doručovania paketov a maximalizovať množstvo doručených paketov [3].

#### ZÁVER

Tento článok popisuje základnú myšlienku využitia MANET v 6G sieti. Tieto siete budú pokračovateľmi sietí štvrtej a piatej generácie. 6G je teda koncept revolučnej mobilnej siete, ktorá bude poskytovať nový typ bezdrôtovej dátovej komunikácie a bude implementovať nové typy smerovacích algoritmov využívajúcich umelú inteligenciu. Základným prvkom 6G sietí budú aj MANET siete, ktoré budú využívať alternatívne smerovacie algoritmy, ktoré budú aplikovať do procesu smerovania strojové učenie, technológiu blockchain, ako aj viacestné smerovacie algoritmy. Cieľom je teda efektívne využívanie dostupného frekvenčného pásma, poskytovať vysoké prenosové rýchlosti, dosahovať nízke oneskorenie a poskytovať nové typy aplikácií. 6G siete majú veľký potenciál pre prepojenie všetkého, čo patrí do bežného života ľudí, ako je zdravotná starostlivosť, životné prostredie, inteligentná domácnosť a podobne.

#### POĎAKOVANIE

Táto publikácia vznikla vďaka podpore Agentúry na podporu výskumu a vývoja v rámci projektu c. APVV-17-0208 - Odolné mobilné siete na doručovanie obsahu.

#### REFERENCIE

- [1] P. Holslin, "What Is 6G Internet and What Will It Look Like? [online]. 2020-08-19. [cit. 2020-11-09]. Dostupné na internete: <<https://www.highspeedinternet.com/resources/6g-internet>>.
- [2] M.A.A. Careem, A. Dutta, "Reputation based Routing in MANET using Blockchain,," International Conference on COMMunication Systems & NETWORKS (COMSNETS), 2020.
- [3] A.K. Das, R. Chaki, K. N. Dey, "Secure energy efficient routing protocol for wireless sensor network,," Foundations of Computing and Decision Sciences, Vol. 41, No. 1, pp. 3–27, 2016.
- [4] K. Sheth, K. Patel, H. Shah, S. Tanwar, R. Gupta, N. Kumar, "A taxonomy of AI techniques for 6G communication networks. 2020.
- [5] Y. Zhou, L. Liu, L. LuWanga, N. Hui, X. Cui, Y. Peng, "Service aware 6G:an intelligent and open network based on convergence of communication, computing and caching. 2020.
- [6] L. Khan, I. Yaqoob, M. Imran, Z. Han, Ch. Hong, "6G Wireless Systems: A Vision Architectural Elements, and Future Directions. 2016.
- [7] H. Chen, J. Li, T. Li, K. Tu, S. Tang, Z. Quing, "6G Wireless Communications: Security Technologies and Research Challenges. 2020.
- [8] N. Nguyen, L. Tran, J. Loven, M. Partala, S. Pirttikangas, "Privacy-aware blockchain innovation for 6G: Challenges and opportunities,," 6G Wireless Summit (6G SUMMIT), pp. 1–5, 2020.
- [9] S. Elmeadawy, R. M. Shubair, "6G wireless communications: Future technologies and research challenges,," International Conference on Electrical and Computing Technologies and Applications (ICECTA), pp. 1–5, 2019.
- [10] T. Zhang, Y. Xu, J. Loo, D. Yang, and L. Xiao, "Joint computation and communication design for uav-assisted obile edge computing in IoT,," IEEE Transactions on Industrial Informatics, vol. 16, no. 8, pp. 5505–5516, 2020.

# Parametre náhradného modelu fotovoltického článku

<sup>1</sup>Tomáš MANIAK, <sup>2</sup>Marek PAVLÍK

<sup>1,2</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>tomas.maniak@student.tuke.sk, <sup>2</sup>marek.pavlik@tuke.sk

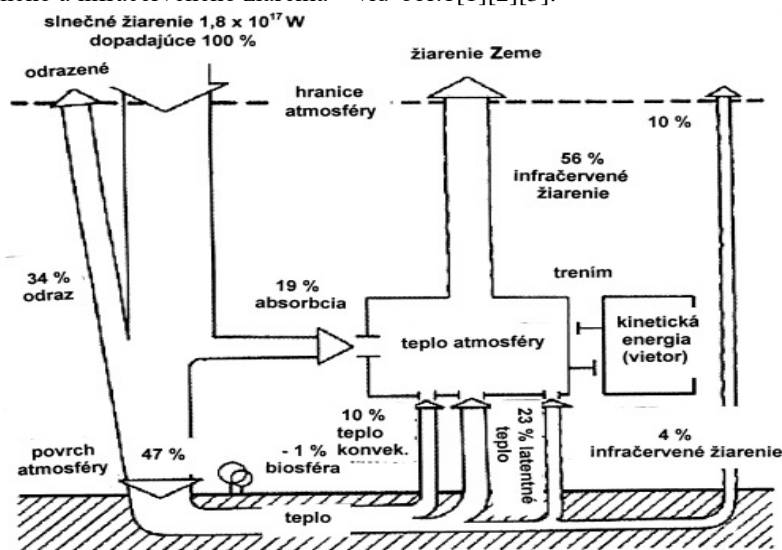
**Abstract** — Tento príspevok približuje ako parametre náhradného modelu fotovoltického (FV) článku ovplyvňujú množstvo vyrobenej elektrickej energie. Je potrebné si uvedomiť, že samotný výkon FV panela nezaručuje potrebné množstvo vyrobenej elektrickej energie. Vstupuje tu niekoľko parametrov a tento príspevok si dáva za cieľ priblížiť tieto parametre a poukázať na ich vplyv na samotnú výrobu elektrickej energie. Náhradným modelom je možné modelovať niektoré parametre FV článku ako napríklad oteplenie FV článku, poprípade vplyv poklesu intenzity osvetlenia na výrobu FV článku.

**Keywords** — fotovoltický panel, parametre FV panelov, slnečné žiarenie

## I. ÚVOD

Energia zo Slnka je hnacím mechanizmom života na Zemi. Zohrieva atmosféru a povrch Zeme, čím vytvára vietor, takisto zohrieva aj oceány, moria, jazera, rieky a pod., zohrievaním vody dochádza k jej odparovaniu a vzniku neustáleho prebiehajúceho kolobehu vody a taktiež rastliny potrebujú pre svoj rast slnečnú energiu. Z dlhodobého hľadiska slnečná energia vytvára aj fosilné paliva. Slnečná energia a z nej pochádzajúce obnoviteľné zdroje energie môžu byť využité na výrobu všetkých foriem energie, ktoré dnes ľudstvo využíva.

Energia slnečného žiarenia dopadajúceho na Zem za jeden rok je 20 tisíc krát väčšia, ako je celosvetová spotreba energie. Výkon Slnka je približne  $4 \cdot 10^{26}$  W, z toho vrchná časť atmosféry prijíma veľmi malé množstvo tejto energie. Takže na povrch atmosféry dopadne výkon približne  $1,85 \cdot 10^{17}$  W. Tento výkon prenáša celé spektrum vlnových dĺžok, avšak atmosféra Zeme pohlcuje röntgenové žiarenie a taktiež zadržiava aj väčšinu ultrafialového žiarenia. Avšak túto pohltenu alebo inak stratenú energiu môžeme zanedbať. Najväčšia časť dopadajúcej energie na Zem je vo forme viditeľného a infračerveného žiarenia – vid' obr.1[1][2][3].



Obr. 1 Slnečné žiarenie, prenikajúce na zemský povrch

Ako je možné vidieť na obr. 1, z celkového žiarenia, ktoré dopadne na povrch atmosféry sa odrazí späť do vesmíru približne 34 %, ďalej 19 % sa absorbuje v atmosfére, čo spôsobuje zahrievanie atmosféry. A nakoniec 47 % dopadne na zemský povrch. Po všetkých uskutočnených energetických procesoch na Zemi a v atmosfére sa teplo vyžaruje späť do okolitého kozmického priestoru v podobe infračerveného žiarenia [2][5].

## II. ZLOŽKY SLNEČNÉHO ŽIARENIA

Slnečné žiarenie sa skladá z týchto zložiek:

- Priame slnečné žiarenie, ktoré nezmenilo svoj smer a dopadá priamo na zemský povrch pri jasnej oblohe. Jeho intenzita závisí od uhla dopadu slnečného žiarenia a znečistenia atmosféry.
- Difúzne slnečné žiarenie, ktoré vzniká pri každej oblačnosti, kde dochádza k rozptýleniu slnečného žiarenia na atómy, molekuly a aerosóly nachádzajúcich sa v atmosfére. Toto žiarenie dopadá na zemský povrch zo všetkých smerov.
- Odrazene slnečné žiarenie je žiarenie, ktoré sa odráža od okolitých budov, ciest vegetácie a podobne. Aj tieto lúče dopadajú na plochu panela. Intenzita odrazeného žiarenia od snehu v zime je podstatne vyššia ako intenzita v lete pri odraze od vegetácie.

**Globálne žiarenie** je súčet priameho a difúzneho žiarenia ktoré dopadajú na zemský povrch.

**Celkové žiarenie** je súčet priameho, difúzneho a odrazeného žiarenia. Používa sa pri technických výpočtoch na určenie intenzity slnečného žiarenia z ožiarenej plochy.

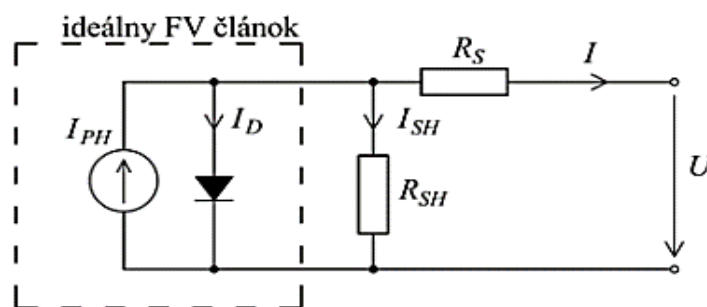
Slnečné žiarenie  $I_C$ , ktoré dopadne na FV panel alebo solárny kolektor je definované vzťahom [2]:

$$I_C = I_P + I_O + I_D \quad (1)$$

kde:  $I_P$  - priame slnečné žiarenie  
 $I_O$  - odrazené slnečné žiarenie  
 $I_D$  - difúzne slnečné žiarenie

## III. PARAMETRE FOTOVOLTICKÝCH ČLÁNKOV

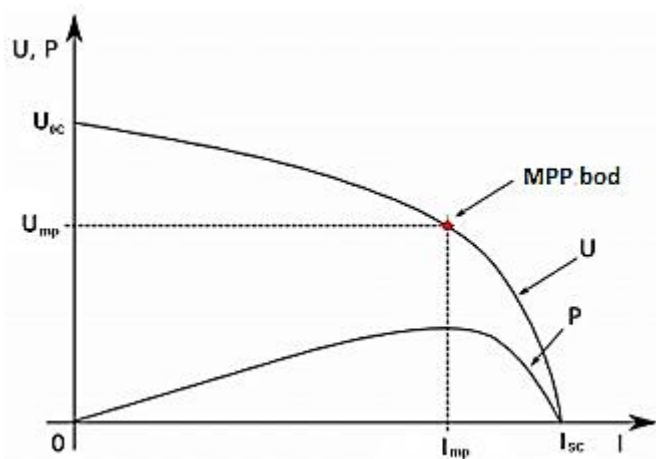
Fotovoltický článok je vo svojej podstate prúdový zdroj, ktorého veľkosť vyrobeného prúdu závisí od množstva dopadajúceho svetla na plochu FV panela. Na obrázku Obr. 2 je zobrazená náhradná schéma FV článku. Tento model obsahuje zdroj prúdu  $I_{PH}$ , jednu diódu a sériový odpor  $R_S$ , ktorý predstavuje odpor materiálu polovodiča a odpor kontaktov. Dióda ma tiež paralelný odpor  $R_{SH}$ , ktorý definuje rozsiahle defekty. V ľavej časti schémy je zobrazený ideálny FV článok (čiarkovanou čiarou ohraničená oblasť), ktorý sa skladá len zo zdroja prúdu a diódy. Avšak v praxi ideálny FV článok neexistuje, pretože sa tam vyskytujú vyššie spomenuté odpory [3][4].



Obr. 2 Náhradná schéma FV článku

FV článok je možné si zjednodušene predstaviť ako PN prieťah v priepustnom smere, ktorý má svoju konkrétnu VA charakteristiku a s tým spojené základné elektrické veličiny určujúce vlastnosti každého modulu. Ideálna VA charakteristika FV modulu má priebeh zobrazený na Obr.3. Priebeh VA charakteristiky je značne závislý od teploty a intenzity slnečného žiarenia [1][3].

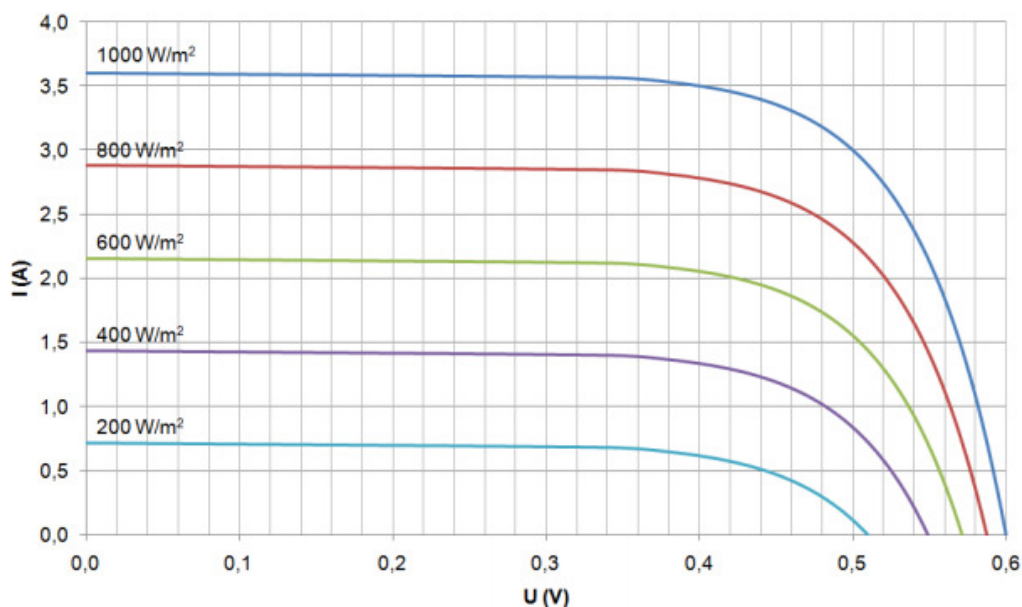
Na náhradnej schéme FV článku boli vytvorené simulácie pre zistenie, ako ovplyvňuje slnečná energia a teplota FV článku samotnú výrobu elektrickej energie. Množstvo vyrobenej elektrickej energie je možné definovať podľa V-A charakteristiky FV článku. Ako už bolo skôr spomenuté, vplyv slnečného žiarenia a vplyv teploty FV článku je možné definovať nastavením hodnotami na sériovom a paralelnom odpore.



Obr. 3 V-A charakteristika FV článku[2]

Najväčší vplyv na výkon FV článkov má intenzita dopadajúceho žiarenia. To je zapríčinené fyzikálnou podstatou premeny slnečnej energie na elektrickú energiu. Čím je viac dopadajúceho slnečného žiarenia na článok, tým viac dokáže FV článok vyrobiť elektrickej energie.

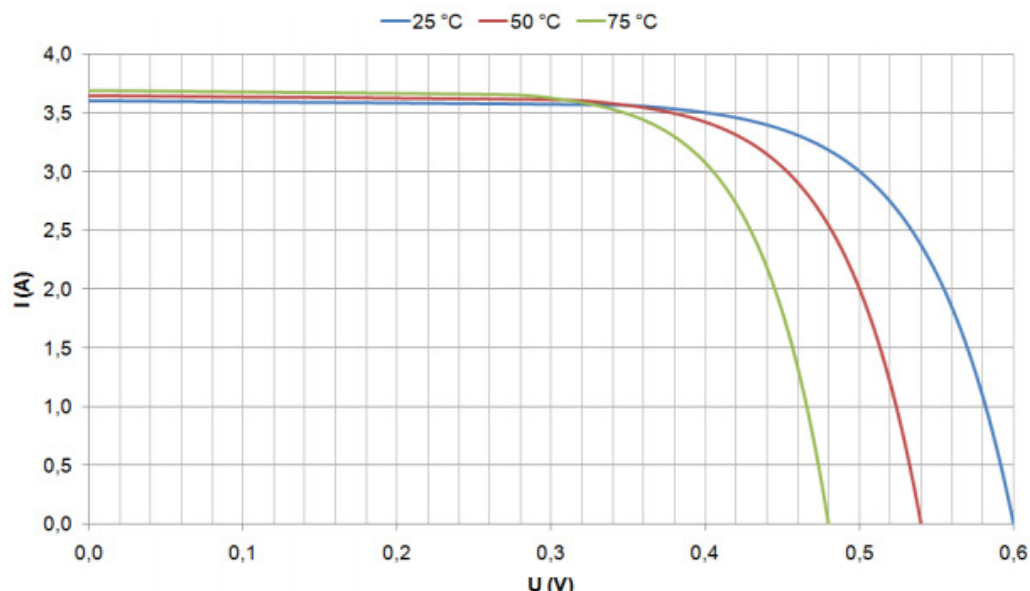
Boli vytvorených 5 simulácií pre rôznu intenzitu slnečného žiarenia. Z výsledkov vyplýva, že intenzita slnečného žiarenia najviac ovplyvňuje prúd nakrátko  $I_{sc}$ , čo je možné vidieť na Obr.4. Pri zvyšovaní intenzity slnečného žiarenia narastá hodnota prúdu nakrátko  $I_{sc}$  omnoho výraznejšie ako hodnota napätia naprázdno  $U_{oc}$ .



Obr. 4 Vplyv intenzity slnečného žiarenia na V-A charakteristiku

Na výkon FV panela má taktiež veľký vplyv teplota – viď Obr.5. Pri dlhšej trvajúcej zvýšenej intenzite žiarenia alebo pri zhoršených chladiacich podmienkach v miestach, kde teplota vzduchu dosahuje až 40 °C, môže dôjsť k nahriatiu článkov až na 80 °C. Pri takto vysokých teplotách dochádza k zmene elektrických vlastností článkov. Tieto zmeny sa najviac prejavujú pri znížení napätia naprázdno  $U_{oc}$ . Na obr.5 je možné vidieť výsledok troch simulácií. Boli vytvorené simulácie, kedy mal FV článok teplotu 25°C, 50°C a 75 °C. Z výsledkov vyplýva, že so zvýšenou teplotou FV článku klesá napätie naprázdno  $U_{oc}$ . Pokles tohto napätia na za následok pokles celkovej výroby z FV článku a teda aj pokles výroby z celého FV panela.

Prevádzkové teploty panela sa pohybujú v rozsahu od – 35 °C do 85 °C. Výkon pri rovnakej intenzite svetelného žiarenia je omnoho väčší v zimnom období ako výkon v letnom období.



Obr. 5 Vplyv teploty na V-A charakteristiku

#### IV. ZÁVER A DISKUSIA

Nie len slnečné žiarenia má významný vplyv na množstvo vyrobenej elektrickej energie z fotovoltických panelov, resp. článkov. Pomerne podstatný vplyv majú aj samotné parametre FV panela. Štítkový údaj o maximálnom výkone je síce podstatný parameter FV panela, avšak je tu viacero ďalších parametrov, ktoré ovplyvňujú množstvo vyrobenej elektrickej energie. V príspevku sa spomínajú aj parametre - sériový odpor  $R_S$ , ktorý predstavuje odpor materiálu polovodiča a odpor kontaktov, ktorý má za následok prehrievanie FV článku. Z výsledkov zobrazených na obr.5 je možné vidieť ako vplyva teplota FV článku na výrobu FV článku.

Dióda ma tiež paralelný odpor  $R_{SH}$ , ktorý je spôsobený rozsiahlymi defektmi. Zmeny týchto odporov významne ovplyvňujú množstvo vyrobenej elektrickej energie. Príspevok je zameraný aj na vplyv slnečného žiarenia a teploty na vyrobenú elektrickú energiu.

Najväčší vplyv na výkon FV článkov ma intenzita dopadajúceho žiarenia. To je zapríčinené fyzikálnou podstatou premeny slnečnej energie na elektrickú energiu, teda čím je viac dopadaného slnečného žiarenia na článok, tým viac dokáže FV článok vyrábať elektrickú energiu. Intenzita slnečného žiarenia najviac ovplyvňuje prúd nakrátko  $I_{SC}$  a to je možné vidieť na Obr.4, kde pri zvyšovaní intenzity narastá hodnota prúdu nakrátko  $I_{SC}$  omnoho výraznejšie ako hodnota napätia naprázdno  $U_{OC}$ .

Na výkon FV panela ma taktiež veľký vplyv teplota – vid' Obr.5. Pri dlhšej trvajúcej zvýšenej intenzite žiarenia alebo pri zhoršených chladiacich podmienkach v miestach, kde teplota vzduchu dosahuje až 40 °C, môže dôjsť k zahriatiu článkov až na 80 °C. Pri takto vysokých teplotách dochádza k zmene elektrických vlastností článkov. Tieto zmeny sa najviac prejavujú pri znížení napätia naprázdno  $U_{OC}$ .

#### ACKNOWLEDGMENT

Tento príspevok vznikol vďaka podpore projektu APVV-19-0576 "Sebestačnosť elektroenergetiky v podmienkach liberalizovaného trhu s elektrinou".

#### REFERENCES

- [1] KOLCUN, Michal et al.: Výskum charakteristík fotovoltických komponentov pre efektívne projektovanie solárnych systémov. 1.vyd. Technická univerzita v Košiciach. 2014. 120 s. ISBN 978-80-553-1961-2.
- [2] KOLCUN, Michal et al.: Elektrárne. 1.vyd. Technická univerzita v Košiciach. 2017. 202 s. ISBN 978-80-553-3119-5.
- [3] MASTNÝ, Peter et al.: Obnoviteľné zdroje elektrickej energie . 1.vyd. České vysoké učení technické v Praze . 2011. 256 s. ISBN 978-80-01-04937-2
- [4] KAPLANI E.; KAPLANIS S. Energy performance and degradation over 20 years performance of BP c-Si PV modules. Simulation Modelling Practice and Theory, Volume 19, Issue 4, April 2011, Pages 1201-1211, ISSN 1569-190X.
- [5] KONTGES , Marc et al.: Review of Failures of Photovoltaic Modules . 1.vyd. IEA-PVPS. 2014. 140 s. ISBN 978-3-906042-16-9



# Analýza vplyvu predspracovania dát na vybrané modely analýzy dát

<sup>1</sup>Zuzana PELLA, <sup>2</sup>Juraj KANUŠČÁK

<sup>1,2</sup> Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>zuzana.pella@tuke.sk, <sup>2</sup>juraj.kanuscak@student.tuke.sk

**Abstrakt** — Cieľom tejto práce je poukázať na dôležitosť fázy predspracovania dát pre klasifikačný aj regresný typ úlohy objavovania znalostí. Za týmto účelom sme sa zamerali na vybrané metódy predspracovania dát (normalizačné techniky, techniky dopĺňania chýbajúcich dát a metódu výberu počtu atribútov) a hodnotíme ich úspešnosť. Získané výsledky nenaznačujú jednoznačnú prevahu tej ktorej metódy, poukazujú však na významnosť použitia jednotlivých typov vybraných techník.

**Kľúčové slová** — CRISP-DM, Chýbajúce hodnoty, Normalizácia, Predspracovanie dát, Výber atribútov

## I. ÚVOD

Procesy spracovania a analýzy dát sú pomerne rozšírené už takmer v každej oblasti ľudského života. Každý výskum prebiehajúci na akýchkoľvek dátach sa riadi určitým postupom. Jedným z najrozšírenejších je metodológia CRISP-DM (Cross Industry Standard Process for Data Mining), ktorá ponúka štandardizovanú postupnosť určitých krokov: *Pochopenie cieľa*, *Pochopenie dát*, *Predspracovanie dát*, *Modelovanie*, *Vyhodnotenie* a *Nasadenie*. Časovo najnáročnejšou sa javí fáza *Predspracovania dát*, ktorá zahŕňa viacero rôznych typov úloh. Podľa [1] a [2], príprava dát predstavuje približne 80% úsilia vynaloženého na spracovanie a analýzu dát. Kompenzáciou tejto časovej náročnosti však môže byť šetrenie úsilia a času v neskorších fázach analýzy dát. Dáta, ktoré vstupujú do tejto fázy môžu mať rôzny charakter, rôzny typ, môžu obsahovať duplicitné alebo irelevantné údaje, nesprávne zadané hodnoty alebo vôbec nezadané hodnoty, prípadne hodnoty mimo typického rozsahu sledovaných javov. Problémy, s ktorými sa táto fáza musí vyrovnáť je viacero. Aj to podčiarkuje samotnú dôležitosť spomínanej fázy. Výsledkom *Predspracovania dát* je súbor dát pripravený na fázu *Modelovania*. To, aké kvalitné výsledky je možné získať, závisí aj od toho, ako kvalitne pripravené dáta vystupujú z fázy *Predspracovania* a zároveň vstupujú do fázy *Modelovania*. Dobrá príprava dát je zásadná pre následné vytváranie spoľahlivých modelov s vysokou mierou presnosti [1], [2], [3]. Tým, že CRISP-DM umožňuje aj spätný prechod medzi samotnými fázami, znamená, že fáza *Predspracovania dát* nie je definitívne ukončená, dokiaľ nedôjde k samotnému nasadeniu riešenia.

## II. TYPICKÉ ÚLOHY PREDSPRACOVANIA DÁT

Fázu *Predspracovania dát* je možné rozdeliť na šesť okruhov úloh:

1. čistenie dát,
2. integrácia dát,
3. transformácia dát,
4. redukcia dát,
5. diskretizácia dát,
6. správa atribútov.

Každý z týchto okruhov v sebe zahŕňa viacero typov metód, ktoré samozrejme závisia aj od zvoleného analytického cieľa. Samotné čistenie dát sa musí vysporiadať s chýbajúcimi hodnotami, duplicitou, resp. redundanciou dát a odľahlými hodnotami. Integrácia dát má za úlohu spojiť, a tým aj obohatiť, dáta pochádzajúce z viacerých zdrojov. Tento typ úloh nie je triviálnym procesom, nakoľko dáta sa častokrát vyznačujú heterogénnosťou, rôznou formou, rôznou schémou uchovávanania, úrovňou abstrakcie či odlišnou vzorkou časovania. Transformácia v sebe zahŕňa úlohy ako sú generalizácia (čiže zovšeobecnenie), normalizácia, agregácia alebo aj

vytváranie nových atribútov. Účelom vykonania týchto procesov je zlepšenie vizualizácie, kompatibility či interpretovateľnosti údajov. Časť z týchto úloh možno označiť aj ako úlohy redukcie dát. Sú nimi napr. agregácia alebo generalizácia. Ďalej k nim môžeme zaradiť aj úlohy redukcie dimenzionality, kompresie alebo redukcie v zmysle zníženia počtu sledovaných inštancií. Samotná diskretizácia je rozdielna pre kontrolované a nekontrolované učenie. Diskretizácia dokáže byť nápomocná pri znížení objemu dát, lepšej vizualizácii a reprezentácii údajov. Posledným typom úloh je správa atribútov, ktoré možno považovať za prostriedok na zníženie dimenzionality, ale aj na nahradenie pôvodných charakteristík menším počtom odvodených. Rovnako ako pri všetkých typoch úloh, aj v tomto prípade je hlavnou cieľom napomôcť pri zvyšovaní presnosti modelov [1], [2].

V tejto práci sa však z dôvodu objemu možných úloh nebude venovať všetkým, ale zameriame sa iba na vybrané. Podľa [4], môžeme úlohy pokrývajúce fázu *Predspracovania dát* rozdeliť do nasledujúcich skupín:

1. extrakcia príznakov a ich portabilita,
2. čistenie dát,
3. redukcia dát.

Uvedeného členenia na skupiny metód sa budeme držať aj v nasledujúcej časti predkladanej práce.

#### A. Extrakcia príznakov a ich portabilita

Nakoľko sa dáta môžu nachádzať v rôznych dátových úložiskách vo variabilných formách a typoch, je potrebné ich najskôr získať a následne aj zjednotiť formou transformácie. Dôvodom môže byť aj možnosť porovnania jednotlivých atribútov, čo pri rôznych typoch dát môže byť ťažšie realizovateľné. Avšak je potrebné rátať s hrozbou určitej straty informácií. Transformáciu možno vykonať viacerými metódami, akými sú napr. diskretizácia, binarizácia alebo latentná sémantická analýza [4], [5].

V istom zmysle môžeme napr. diskretizáciu považovať aj za redukčnú metódu, nakoľko numerický typ dát značného rozsahu dokáže transformovať na kategorickú reprezentáciu. Podstatné zastúpenie diskretizácie medzi metódami prípravy dát podčiarkuje aj fakt, že pri spojitých premenných je tento typ úpravy dát takmer nevyhnutnosťou. Ak sa to neudeje implicitne vo fáze predspracovania, tento krok je vykonaný napr. pri vytváraní rozhodovacieho stromu vo fáze modelovania [6].

#### B. Čistenie dát

Čistenie dát je potrebné na zabezpečenie jasnosti a užitočnosti procesu objavovania znalostí. Dáta častokrát obsahujú chyby v dôsledku ľudskej nepozornosti, nedokonalosti softvéru, ktorý s dátami pracuje, alebo chyby prenosu údajov. Okrem toho nie je ničím výnimočným, že dáta sú nekompletné a obsahujú chýbajúce údaje – buď vôbec neexistujú alebo boli z akéhosi dôvodu zamlčané. Čistenie dát môže zahŕňať úlohy ako sú integrácia dát, normalizácia dát, či ošetrovanie chýbajúcich hodnôt.

Integrácia zahŕňa úlohy pre zjednotenie dát pochádzajúcich z viacerých zdrojov [7].

Existuje viacero prístupov ako sa vysporiadať s chýbajúcimi hodnotami. Najtriviálnejšie sú na základe odstránenia záznamov alebo atribútov obsahujúcich chýbajúce hodnoty. Namiesto odstránenia sa môže objaviť snaha chýbajúce hodnoty doplniť napr. pomocou priemernej, či strednej hodnoty atribútu. Viac sofistikovanejšími sú metódy založené na strojovom učení. Takými metódami môžu byť napr. doplnenie na základe algoritmu kNN (tzv. kNN Imputation – kNNI) alebo Expectation-Maximization (EM) [4], [8].

Využitie metód normalizácie, resp. štandardizácie rieši skutočnosť, že atribúty sa častokrát objavujú s odlišnými rozsahmi hodnôt a tým sa stávajú navzájom neporovnateľné. Tento typ metód sa odporúča vykonávať až po rozdelení dátovej množiny na tréningovú a testovaciu z dôvodu možnej straty informácií. Tieto metódy sú nezriedkavo využívané algoritmami, ktoré pracujú na základe vzdialenosti medzi sledovanými dátami [4].

#### C. Redukcia a transformácia dát

Metódy redukcie a transformácie sa využívajú s cieľom znížiť rozsah dát. Tento krok je častokrát realizovaný aj z dôvodu zníženia výpočtovej náročnosti algoritmov, čo teda urýchľuje čas vytvárania modelov. Navyše nie je zriedkavé, že takýto krok vedie k vyššej presnosti modelov [6].

Na redukciu dát je možné sa pozerieť z dvoch hľadísk. Prvé je zamerané na redukciu počtu sledovaných charakteristík (z ang. Feature Selection). Druhé hľadisko sa zameriava na zníženie

počtu sledovaných záznamov – označované ako vzorkovanie (z ang. Sampling). Využitie týchto metód sa radí medzi najpoužívanejšie z fáze predspracovania dát v oblasti kardiológie [7]. Okrem toho, že redukcia veľkosti dát môže viesť k lepšej vizualizácii [6], dochádza tu napríklad aj k odstráneniu charakteristík, ktoré sú irelevantné alebo redundantné [7].

### III. ANALÝZA SÚČASNÉHO STAVU

Podobne ako predchádzajúcu kapitolu, aj túto si rozdelíme podľa pozorovanej metódy predspracovania dát. Najskôr sa zameriame na riešenie problematiky chýbajúcich hodnôt, následne na normalizáciu a uzavrieme to výberom atribútov.

Benhar a kol. [7] vykonali prieskum v oblasti dopĺňania chýbajúcich hodnôt pomocou metód strojového učenia. Výsledkom prieskumu bolo, že zhruba 30% výskumov sa zameriavalo na dopĺňanie na základe hodnotami aritmetického priemeru, konštantnej hodnoty alebo využitím algoritmu Expectation Maximization Imoutation (EM).

Štúdia [6] sa zamerala na porovnanie doplnenia chýbajúcich hodnôt pomocou EM a kNNI. Pôvodný dátový súbor neobsahoval žiadne chýbajúce údaje, preto ich tím autorov vytvoril v miere 10%. Grafické porovnanie súborov po doplnení chýbajúcich hodnôt preukázalo, že hodnoty doplnené pomocou kNNI boli podobnejšie pôvodnej dátovej množine ako hodnoty doplnené pomocou EM.

Pre porovnanie normalizačných techník, sme sa rozhodli preskúmať porovnanie, ktoré vykonal kolektív autorov okolo Jaina [9]. Autori normalizovali 48 dátových súborov pomocou min-max normalizácie a štandardizácie. Na porovnanie úspešnosti normalizačných techník si zvolili Gaussian Kernel ELM klasifikátor. Výsledok porovnania neuprednostnil ani jednu zo zvolených techník, nakoľko v 24 prípadoch sa javila úspešnejšou min-max normalizácia a v 24 prípadoch štandardizácia.

Benhar a kol. [7] hodnotili v svojej práci viacero prístupov pre prípravu dát, okrem už uvedeného porovnanie dopĺňania chýbajúcich hodnôt pozorovali aj vplyv výberu atribútov na výsledok modelovania. Výsledkom porovnania úspešnosti klasifikácie s využitím a bez využitia metód výberu atribútov dospeli v 66% porovnaní k zhodnému výsledku pre oba prístupy. V 30% prípadoch zaznamenali autori zlepšenie výsledku po aplikácii metódy na výber atribútov a v 4% došlo k zhoršeniu výsledkov klasifikácie.

Vo viacerých prípadoch nebolo možné jednoznačne určiť, či, resp. ktorá metóda predspracovania dát mala výraznejší pozitívny vplyv na výsledky vytvorených modelov. Rozhodli sme sa preto vykonať v menšom rozsahu aj vlastné porovnanie vplyvu nasledujúcich metód: normalizácia (min-max normalizácia a štandardizácia), doplnenie chýbajúcich hodnôt (expectation-maximalization a kNNI) a výber atribútov (backwards feature selection).

### IV. NÁVRH, REALIZÁCIA A VYHODNOTENIE EXPERIMENTOV

Pre uskutočnenie experimentov sme si zvolili dve dátové množiny, ktoré sú voľne dostupné na UCI Machine Learning Repository. Všetky experimenty boli vykonané v prostredí RStudio v programovacom jazyku R.

Prvou dátovou množinou bola Z-Alizadeh Sani Data Set, ktorý neobsahoval žiadne chýbajúce hodnoty, dáta boli numerického, nominálneho aj binárneho typu. Dáta popisujú pacientov z hľadiska (ne)prítomnosti ischemickej choroby srdca. Po úprave obsahoval dátový súbor 23 numerických atribútov a 33 atribútov typu *factor*. Dáta boli rozdelené na tréningovú a testovaciu množinu v pomere 70:30.

Druhá dátová množina sa zameriava na lesné požiare v oblasti Portugalska, pričom cieľový atribút popisov rozsah plochy zasiahnutej požiarov. Dáta neobsahovali žiadne chýbajúce hodnoty a charakter atribútov bol numerický alebo nominálny. Pre realizáciu experimentov bolo všetkých 13 atribútov prevedených na dátový typ *numeric*. Rovnako aj tu sme pristúpili k rozdeleniu dát na tréningovú a testovaciu množinu v pomere 70:30.

Oba dátové sety budú použité pre porovnanie vplyvu metód normalizácie, dopĺňania chýbajúcich hodnôt a výberu atribútov. Porovnanie úspešnosti bude realizované pomocou dvoch typov metód strojového učenia: klasifikačný algoritmus *cree* na množine Z-Alizadeh Sani Data Set, lineárna regresia na množine Forest Fire Data Set. Pre klasifikačnú úlohu budeme sledovať metriky presnosti (*accuracy*), senzitivity a špecificity. V prípade lineárnej regresie budeme sledovať metriky presnosti (*accuracy*), MAPE (mean absolute percental error), MAE (mean absolute error), RMSE (root mean squared error), MSE (mean squared error).

#### A. Experimenty na Z-Alizadeh Sani Data Set

Prvú skupinu porovnaní tvorili metódy normalizácie. Algoritmus *cree* nebol nijak špeciálne nastavovaný a najskôr bol aplikovaný na dátovú množinu bez normalizácie. Následne boli

numerické dáta pôvodného dátového súboru normalizované technikou štandardizácie a ako poslednou bola upravená množina pomocou min-max normalizácie. Výsledky aplikácie algoritmu ctree na troch predpripravených dátových setoch uvádzame v nasledujúcej tabuľke:

Tabuľka 1  
 Úspešnosť normalizačných techník pre klasifikačnú úlohu

Typ normalizačnej techniky	Presnosť	Senzitivita	Špecifická
bez normalizačnej techniky	0,798	0,841	0,667
štandardizácia	0,786	0,839	0,636
min-max normalizácia	<b>0,833</b>	<b>0,897</b>	<b>0,692</b>

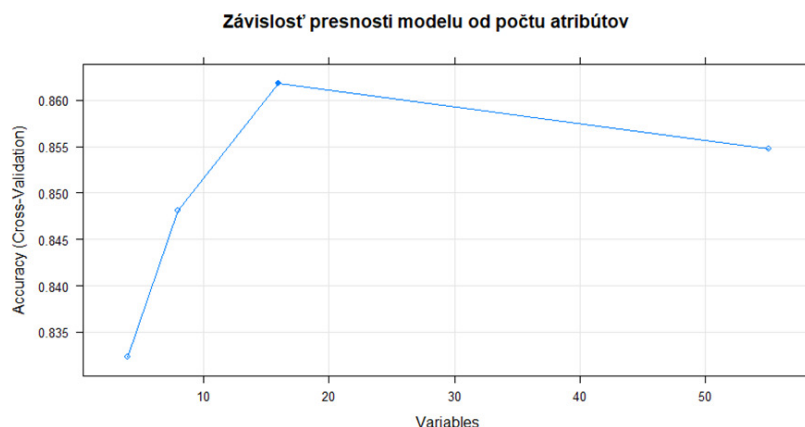
Na základe tabuľky je zjavné, že v prípade popisovaného dátového setu, na ktorý bol aplikovaný algoritmus rozhodovacieho stromu ctree, sa ako lepšia javí technika min-max normalizácie.

V rámci druhého porovnania metód sme sa zamerali na vplyv techník dopĺňania chýbajúcich hodnôt. Najskôr sme umelo vytvorili chýbajúce hodnoty v pomere 5, 10 a 15% numerických dát súboru. Následne sme vykonali doplnenie chýbajúcich hodnôt pomocou algoritmu EM a kNNI. Výsledné porovnanie úspešnosti jednotlivých metód ponúkame v Tabuľke 2.

Tabuľka 2  
 Úspešnosť techník dopĺňania chýbajúcich hodnôt pre klasifikačnú techniku

Úprava dátového setu	Presnosť	Senzitivita	Špecifická
5% chýbajúcich hodnôt v numerických atribútoch			
pôvodný dataset	0,798	0,841	0,667
doplnenie pomocou kNNI	0,798	0,841	0,667
doplnenie pomocou EM	0,798	0,841	0,667
10% chýbajúcich hodnôt v numerických atribútoch			
pôvodný dataset	0,798	0,841	0,667
doplnenie pomocou kNNI	<b>0,810</b>	0,833	<b>0,722</b>
doplnenie pomocou EM	<b>0,810</b>	<b>0,844</b>	0,700
15% chýbajúcich hodnôt v numerických atribútoch			
pôvodný dataset	0,798	<b>0,841</b>	0,667
doplnenie pomocou kNNI	NA	NA	NA
doplnenie pomocou EM	<b>0,810</b>	0,814	<b>0,788</b>

Výsledky v uvedenej tabuľke naznačujú, že použitie techniky na doplnenie chýbajúcich dát dokáže dosiahnuť rovnakú, ak nie dokonca lepšiu presnosť po aplikácii algoritmu ctree. Pri 15% chýbajúcich hodnôt algoritmus kNNI nedokázal doplniť chýbajúce hodnoty, čo si vysvetľujeme nedostatočným počtom dostupných hodnôt pre výpočet dopĺňania. Na základe tejto skutočnosti hodnotíme metódu EM ako úspešnejšiu oproti kNNI.



Obrázok 1. Závislosť presnosti modelu od počtu atribútov

Posledné porovnanie zameriavame na redukcii výberu atribútov. Na Z-Alizadeh Sani Data Set sme aplikovali metódu Backward feature selection, ktorá pre nájdenie optimálneho počtu atribútov využíva klasifikačný algoritmus random forest a 10-násobnú krížovú validáciu. Ako

znázorňuje Obrázok 1, pre dátovú množinu Z-Alizadeh Sani Data Set je najvhodnejšie použiť 16 atribútov, kedy model random forest dosahuje presnosť 86,2%.

### B. Experimenty na Forest Fire Data Set

Postupnosť experimentov normalizácie pre dátový set Forest Fire Data Set odpovedá postupnosti experimentov pre dátový súbor Z-Alizadeh Sani Data Set. Nakoľko normalizácia dát nám znemožňuje použiť všetky zamýšľané metriky porovnania úspešnosti pre lineárnu regresiu, výsledok sme porovnávali len pomocou metriky presnosti a metriky MAPE.

Tabuľka 3  
 Úspešnosť normalizačných techník pre regresný typ úlohy

Typ normalizačnej techniky	Presnosť	MAPE
bez normalizačnej techniky	<b>98,391</b>	<b>1,609</b>
štandardizácia	79,052	20,948
min-max normalizácia	87,269	12,731

V prípade regresnej úlohy nám normalizačné techniky znížili presnosť modelu. Ak by sme však mali vybrať, ktorá poskytuje lepší výsledok, tak z hľadiska presnosti aj priemernej absolútnej percentuálnej chyby (MAPE) by to bola min-max normalizácia.

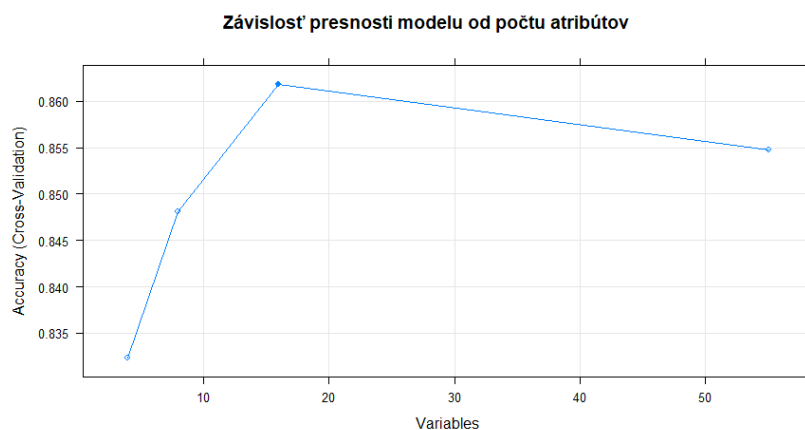
Nakoľko tento dátový súbor predstavuje ďaleko objemnejšie dáta ako prvý dátový set, pri vytváraní chýbajúcich hodnôt sme vyskúšali vytvoriť 10, 15 a 20% chýbajúcich hodnôt. Pre algoritmus kNNI sme nastavili  $k = 5$ . Získané výsledky experimentov uvádzame v Tabuľke 5.

Tabuľka 4  
 Úspešnosť techník dopĺňania chýbajúcich hodnôt pre regresný typ úlohy

Úprava dátového setu	Presnosť	MAPE	MAE	RMSE	MSE
10% chýbajúcich hodnôt v numerických atribútoch					
pôvodný dataset	<b>98,594</b>	<b>1,406</b>	21,561	<b>71,187</b>	<b>5067,626</b>
doplnenie pomocou kNNI	97,914	2,086	<b>21,133</b>	71,197	5069,061
doplnenie pomocou EM	98,366	1,634	29,031	75,191	5653,712
15% chýbajúcich hodnôt v numerických atribútoch					
pôvodný dataset	<b>98,594</b>	<b>1,406</b>	21,561	71,187	5067,626
doplnenie pomocou kNNI	95,775	4,225	<b>21,037</b>	<b>70,455</b>	<b>4963,894</b>
doplnenie pomocou EM	98,535	1,464	30,508	74,282	5517,818
20% chýbajúcich hodnôt v numerických atribútoch					
pôvodný dataset	<b>98,594</b>	<b>1,406</b>	21,561	71,187	5067,626
doplnenie pomocou kNNI	97,529	2,471	<b>18,922</b>	68,173	4721,502
doplnenie pomocou EM	98,518	1,482	37,611	<b>66,013</b>	<b>4357,686</b>

Z hľadiska presnosti žiaden typ doplnenia chýbajúcich hodnôt nedosahuje takú dobrú hodnotu ako pôvodný dátový set. Pri samotnom porovnaní metód dopĺňania chýbajúcich hodnôt sa lepšou javí metóda EM pri 10,15 aj 20% chýbajúcich hodnôt. Metrika MAPE preukazuje to isté, čo presnosť vytvorených modelov. Hodnota MAE je vo všetkých prípadoch najlepšia pri doplnení pomocou kNNI. Metriky RMSE a MSE pre rozdielne množstvá chýbajúcich hodnôt dosiahli rozdielne výsledky. Ak ale máme vybrať iba medzi metódami dopĺňania kNNI a EM, tak pri 10 a 15% chýbajúcich hodnôt sa lepšia javí metóda kNNI, pri 20% chýbajúcich hodnôt sa do popredia dostáva metóda EM.

Pri výbere atribútov pomocou metódy Backwards feature selection na skúmanom dátovom sete sme sledovali metriky RMSE a MAE. Obe metriky mali pri zvyšujúcom sa počte atribútov tendenciu stúpať. Závislosť hodnoty RMSE od počtu atribútov je znázornená na Obrázku 2. Využitie metódy redukcie počtu atribútov sa aj v tomto prípade javilo ako užitočné.



Obrázok 2. Závislosť metriky RMSE od počtu atribútov

## V. ZÁVER

Na základe vykonaných experimentov nie je možné jednoznačne určiť najvhodnejší typ úlohy predspracovania dát. Taktiež na základe porovnania jednotlivých typov metód nevieme vybrať jednu najlepšiu. V prípade prvej skupiny experimentov sa pre klasifikačný typ úlohy ukazovala lepšou min-max normalizácia, pri regresnom type úlohy však presnosť modelu po normalizácii klesla, no pri samotnom porovnaní normalizačných techník dosiahla lepšie výsledky min-max normalizácia. Experimenty zamerané na chýbajúce hodnoty preukazujú svoj značný význam najmä pri klasifikačnom type úlohy. Regresný typ úlohy si z hľadiska presnosti dopĺňanie chýbajúcich hodnôt až tak nevyžaduje, no metriky ako MAE, RMSE a MSE sa prikláňajú na stranu kNNI. Výber počtu atribútov (teda redukcia počtu atribútov) sa javí významným aj pri menšom počte atribútov pôvodného dátového setu. Pri oboch typoch úloh, klasifikačnej aj regresnej, zníženie počtu atribútov zvýšilo presnosť modelov. Na záver však hodnotíme, že neexistuje žiaden univerzálny postup pre predspracovanie každej dátovej množiny v rámci každého typu úloh. Výber metód predspracovania dát je silne aplikačne závislý a dobré pochopenie dát s ohľadom na cieľ úlohy dolovania dát dokáže pomôcť pri správnom výbere metód predspracovania dát.

## POĎAKOVANIE

Táto práca vznikla vďaka podpore Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-17-0550 a Vedeckej grantovej agentúry MŠVVaŠ SR a SAV projekt č. 1/0685/21.

## REFERENCIE

- [1] S. Zhang, C. Zhang, Q. Yang, *Data Preparation for Data Mining*. Applied Artificial Intelligence, Vol. 17(5-6), 2003, pp. 375-381.
- [2] Z. S. Abdallah, L. Du, G. I. Webb, *Data Preparation*. Encyclopedia of Machine Learning and Data Mining, (Editors: C. Sammut, G. I. Webb), Boston: Springer, 2017.
- [3] S. García, J. Luengo, F. Herrera, *Data Preprocessing in Data Mining*, Switzerland: Springer, 2015.
- [4] J. Paralič, *Objavovanie znalosti v databázach*. Košice, 2003.
- [5] Ch. C. Aggarwal, *Data Mining: The Textbook*. New York: Springer, 2015.
- [6] S. García, J. Luengo, F. Herrera, *Tutorial on practical tips of the most influential data preprocessing algorithms in data mining*. Knowledge-Based Systems, Vol. 98, 2016, pp. 1-29.
- [7] H. Benhar, A. Idri, J.L. Fernández-Alemán, *Data preprocessing for heart disease classification: A systematic literature review*. Computer Methods and Programs in Biomedicine, Vol. 195, 2020, pp. 30.
- [8] J. R. Hutchinson, *Axisymmetric flexural vibrations of a thick free circular plate*. Journal of Applied Mechanics, Vol. 46(1), 1979, pp. 139-144.
- [9] S. Jain, S. Shukla, R. Wadhvani, *Dynamic selection of normalization techniques using data complexity measured*. Expert Systems with Applications, Vol. 160, 2018, pp. 252-262.

# Modelovanie účinnosti tienenia elektromagnetického poľa

<sup>1</sup>Tomáš ONDRUŠKO, <sup>2</sup>Marek PAVLÍK

<sup>1,2</sup> Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

<sup>1</sup>tomas.ondrusko@student.tuke.sk, <sup>2</sup>marek.pavlik@tuke.sk

**Abstract** — Tento príspevok je zameraný na vytvorenie modelu pre účely určenia účinnosti tienenia. Následne boli skúmané materiály s rôznymi vlastnosťami. Z výsledkom modelovania v programe Ansys boli určené krivky účinnosti tienenia v sledovanom frekvenčnom rozsahu od 2GHz do 8GHz. Sledované materiály boli sklo, plexisklo, mramor, tvrdá guma a ferit.

**Keywords** — účinnosť tienenia elektromagnetického poľa, žiarenie, bariéra

## I. ÚVOD

Pribúdajúcimi rokmi sme svedkami nepozorovateľného rozširovania a nárastu technických zdrojov elektromagnetického poľa a tým aj rozvojom bezdrôtových informačných technológií. Veľké haló je momentálne v súvislosti s novovzniknutou sieťou 5G, ktorá je v súčasnosti veľmi podporovaná. Na jednej strane je to pre náš veľký prínos no na strane druhej nárast elektromagnetických polí prináša aj isté negatíva[1][2].

Dennodenne sme v kontakte s elektrickými prístrojmi, ktoré generujú elektromagnetické polia (EMP). Preto je dobré vedieť o ich škodlivých účinkoch na náš organizmus a taktiež vedieť ako sa pred takýmito elektromagnetickými poľami dá chrániť[1][3].

Práve kvôli takýmto dôvodom je dobré aby sa elektromagnetické polia skúmali. Najbežnejším spôsobom ich skúmania je práve modelovanie takýchto polí použitím rôznych počítačových programov, ktoré sú schopné simulovať prestup takéhoto poľa rôznymi materiálmi a následne ich dokázu spoľahlivo analyzovať[1][3].

Elektromagnetické pole je pole tvorené dvomi zložkami, elektrickou a magnetickou zložkou. Elektrické pole je všade tam, kde sa nachádza nejaká forma elektrickej energie. Vzniká prítomnosťou elektricky nabitých častíc v elektrických zariadeniach, ktoré sú pripojené na zdroj elektrickej energie, aj keď nimi práve nepreteká prúd. Magnetické pole je charakteristické tým, že vznikne iba tam, kde je prítomný prechod elektrického prúdu. Elektrické pole je definované nasledovne: [1][3]

- vektorom elektrickej indukcie  $D$  s jednotkou  $A \cdot s/m^2$
- vektorom intenzity elektrického poľa  $E$  s jednotkou  $V/m$

Ako už bolo spomenuté magnetické pole vzniká pohybom nabitých častíc, presnejšie pohybom elektrónov v konduktore, ako príklad môžeme uviesť obyčajný kábel. Pohybom týchto elektrónov káblom vzniká elektrický prúd. Čím je tento prúd väčší, tým je aj veľkosť magnetického poľa väčšia, no zvyšovaním vzdialenosti od zdroja energie jeho veľkosť klesá. Je podobne ako pole elektrické definované dvoma vektormi: [1][3]

- vektor magnetickej indukcie  $B$  s jednotkou  $T$
- vektor intenzity magnetického poľa  $H$  s jednotkou  $A/m$

Vzájomným spojením týchto dvoch polí, t.j. elektrického a magnetického dostaneme pole elektromagnetické a jej zložky sú na seba navzájom kolmé. Z hľadiska frekvencie obsahuje všetky zložky elektromagnetickej energie a to rádiové vlny, mikrovlny, infračervené žiarenie, viditeľné svetlo, ultrafialové žiarenie, röntgenové svetlo a  $\gamma$ -žiarenie[1][3].

Elektromagnetické pole je špecifické tým, že sa v priestore pohybuje vo forme vln. V niektorých prípadoch však môže mať formu energie vytvorenej z častíc, tzv. fotónov. Fotón je častica, ktorá zahŕňa energiu, ktorá je v podstate základnou jednotkou elektromagnetickej vlny a závisí od frekvencie a vlnovej dĺžky. To znamená, že čím je frekvencia vyššia/vlnová dĺžka kratšia, tým väčšia energia je nahromadená vo fotóne. Táto energia sa meria v elektronvoltoch

(eV) [1][3].

Fotóny, ako elementárne častice vytvárajú röntgenové a  $\gamma$ -žiarenie obsahujúce veľké množstvo energie, ale iba ak sa bavíme o vysokofrekvenčných poliach. V poliach s nízkou frekvenciou množstvo energie radikálne klesá aj o niekoľko desiatok jednotiek. Vznikajú pri procesoch ionizácie, kedy sa elektróny odtrhnú od atómu resp. molekuly a vyžiari sa fotón. Na atómoch vznikajú rôzne genetické poškodenia a deformácie, ktoré môžeme zhrnúť ako biologické poškodenia. K týmto procesom sú potrebné fotóny s vysokým obsahom energie. Sem môžeme zaradiť práve fotóny röntgenového žiarenia a fotóny  $\gamma$ -žiarenia. Jedna takáto kvantová reakcia je schopná spôsobiť ionizáciu a už spomínané biologické poškodenia, v dôsledku veľkej energie fotónu, ktorá presahuje minimálnu hodnotu potrebnú k ionizácii a to je 10 eV. Práve preto sú práve röntgenové a  $\gamma$ -žiarenie príkladmi ionizujúceho žiarenia, ktoré spĺňajú toto kritérium. Efekt, ktorý sa využíva pri generovaní jadrovej energie (často označovaná iba ako radiácia) sa nazýva ionizácia [1][3].

## II. MODELOVANIE ELEKTROMAGNETICKÉHO POĽA

Pojmom model môžeme označiť materiálnu, ale aj nemateriálnu napodobeninu objektu, pričom nemusíme poznať účel jeho použitia. K ich vytvoreniu sa väčšinou pristupuje vtedy, keď je skúmanie javov v ich prirodzenej forme ťažko realizovateľné alebo sa nedá zrealizovať vôbec.

Na začiatku každého problému, ktorý sa rozhodneme riešiť, je potrebné získať predbežné vedomosti o daných procesoch, javoch a objektoch. Niekedy je táto informácia tak dôležitá a bohatá, že problém je možné riešiť bez toho aby sa pristupovalo k ďalšiemu skúmaniu objektu, no k takejto informácii je aj nevyhnutnosťou vytvoriť adekvátny model. Ak sa nám takúto informáciu nepodarí získať, musíme pokračovať v skúmaní daného objektu. V niektorých prípadoch sa môže stať že to možné nebude.

Vo fyzike sme sa stretli s elementárnymi metódami výpočtu polí. Ich použitie je síce jednoduché, ale dokážeme ho použiť iba pri jednoduchom usporiadaní elektród v tvare guľ alebo valcov. Analytické riešenie je neľahké, z dôvodu riešenia Laplaceovej alebo Poissonovej diferenciálnej rovnice. Taktiež jeho použitie je veľmi obmedzené podobne ako elementárne metódy. Výpočet poľa je možné aj použitím integrálnych rovníc. Riešenie diferenciálnych a integrálnych rovníc je veľmi náročné a v kombinácii s problémami z praxe prakticky nemožné, no ich numerické riešenie na rozdiel od predchádzajúcich dvoch je jednoduché, pretože je možné použiť výpočtovú techniku. Aj napriek tomu majú elementárne a analytické metódy nezastupiteľný význam, najmä pri testovaní a kontrole výstupov z programov založených na numerickom riešení [1][3].

### *Elementárne metódy*

- Gaussova veta elektrostatiky
- princíp zrkadlenia, superpozícia
- výpočet poľa pomocou Coulombovho zákona

### *Analytické metódy*

- priama integrácia, iba pre 1D problém
- separácia premenných, teoreticky neobmedzené použitie, ale pre ľubovoľné usporiadanie elektród veľmi náročné

### *Numerické metódy*

- momentová metóda (riešenie diferenciálnej, integrálnej rovnice)
- metóda hraničných prvkov (riešenie integrálnej rovnice)
- metóda indukovaných nábojov (riešenie integrálnej rovnice)
- metóda konečných objemov (riešenie diferenciálnej rovnice)
- metóda konečných diferencií (riešenie diferenciálnej rovnice)
- metóda konečných prvkov (riešenie diferenciálnej rovnice)

Ich spoločným znakom je, že sa nahradí presné riešenie diferenciálnej alebo integrálnej rovnice riešením približným a to zadaním nasledujúcich parametrov:

- rozmerov a tvaru usporiadania, pričom oblasť v ktorej sa hľadá riešenie sa rozdelí na menšie oblasti (tzv. podoblasti), tie sa ďalej delia na elementy a uzly s vlastnými súradnicami
- materiálových vlastností prostredia jednotlivých podoblastí
- rozloženia zdrojov poľa a ich vlastností v jednotlivých podoblastiach a plochách
- diferenciálnej alebo integrálnej rovnice poľa
- okrajových podmienok na hraniciach podoblastí
- ďalších vzťahov, odvodených pre výpočet sekundárnych veličín

### *Inverzné a optimalizačné metódy*

V poslednej dobe sa rozvíjajú najmä vďaka numerickým metódam. Cieľom inverzných metód je nájsť k pôvodnému rozloženiu poľa v obmedzenej oblasti rozloženie zdrojov mimo túto oblasť



a k nameraným hodnotám poľa nájsť možné rozloženie zdrojov, alebo materiálových vlastností skúmaného prostredia. Optimalizačné metódy zase slúžia napríklad aby maximálna intenzita v izolácii bola čo najnižšia, rozmery zariadenia, váha alebo cena boli čo najnižšie atď.[10]

### III. ÚČINNOSŤ ELEKTROMAGNETICKÉHO POĽA

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 pričku. Lepším ukazovateľom tienenia je účinnosť tienenia  $SE$ . Účinnosť tienenia je možné charakterizovať ako súčet odrazu elektromagnetického poľa  $R$  a absorpcie elektromagnetického poľa  $A$ : [4][5]

$$SE = A + R \quad (1)$$

Koeficient absorpcie sa určí podľa vzťahu (2) ako: [4][5]

$$A = 8,69 \frac{t}{\delta} \quad (2)$$

kde  $t$  je hrúbka tieniaceho materiálu a  $\delta$  je hĺbka vniku elektromagnetického poľa do materiálu. Hĺbka vniku sa vypočíta podľa vzťahu (3) ako: [4][5]

$$\delta = \sqrt{\frac{2}{\omega \cdot \mu \cdot \sigma}} \quad (3)$$

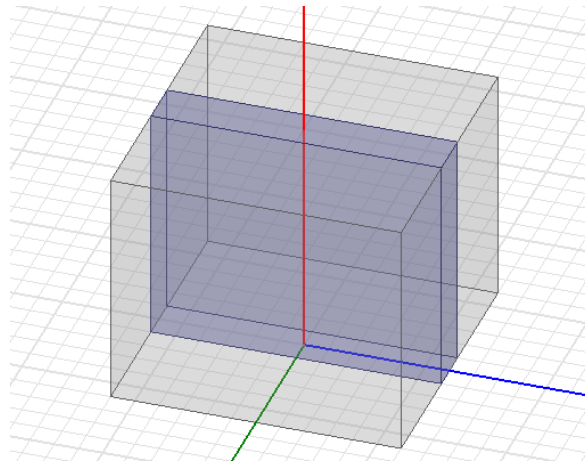
kde  $\mu$  je permeabilita, v ktorej je zahrnutá aj permeabilita tieniaceho materiálu,  $\sigma$  je merná vodivosť tieniaceho materiálu.

Koeficient odrazu  $R$  sa určí podľa vzťahu (4) ako: [4][5]

$$R = 20 \cdot \log \left( \frac{1}{4} \sqrt{\frac{\sigma}{\omega \cdot \mu_r \cdot \epsilon_0}} \right) \quad (4)$$

kde  $\mu_r$  je permeabilita tieniaceho materiálu,  $\epsilon_0$  je permitivita vákua.

Pre účely určenia účinnosti tienenia elektromagnetického poľa bol vytvorený model v programe Ansys – Vid' Obr.1.

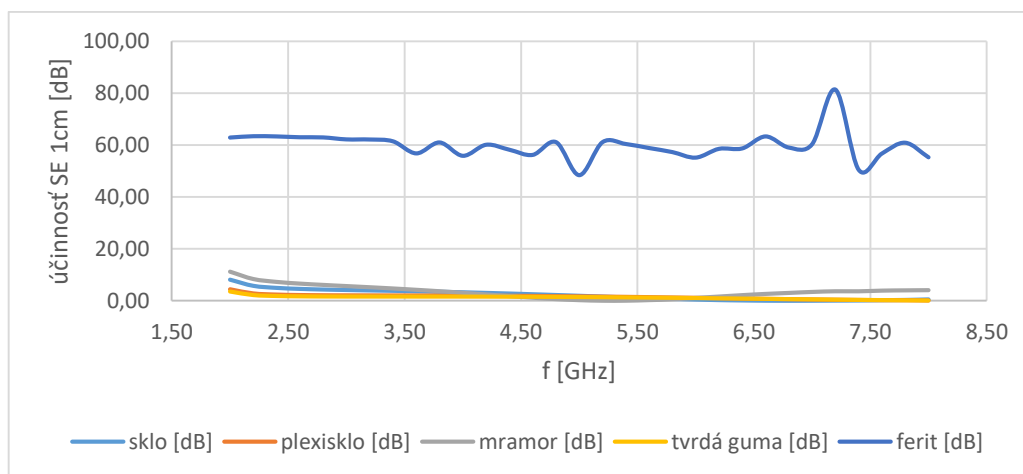


Obr. 1 Navrhnutý model pre účely výpočtu účinnosti tienenia elektromagnetického poľa

Model bol vytvorený pre 5 rôznych materiálov a to sklo, plexisklo, mramor, tvrdá guma a ferit. Vlastnosti materiálov je možné vidieť v Tab.1. Tienenie bolo stanovené s hrúbkou 1 cm. Účinnosť tienenia bola riešená vo frekvenčnej oblasti od 2 GHz do 8GHz. Na Obr. 2 je možné vidieť priebeh účinnosti tienenia elektromagnetického poľa. Z výsledkov vyplýva, že v sledovanom frekvenčnom rozsahu dosahuje najvyššie hodnoty účinnosti tienenia dosahuje materiál ferit, čím bol potvrdený predpoklad, ktorý bol zrejmy na základe vlastností akými sú relatívna permitivita a relatívna permeabilita. Aj tieto vlastnosti ovplyvňujú schopnosť materiálu brániť prieniku elektromagnetického poľa. Zahraničné publikácie popisujú, že materiálové konštanty sa menia so sledovanou frekvenciou. V tomto prípade bol model vytvorený tak, že materiálovú konštanty sa s frekvenciou nemenili a teda výsledky sa môžu líšiť vzhľadom na hodnoty, ktoré boli merané v bezodrazovej komore.

Tabuľka 1  
 Vlastnosti materiálov

Použité materiály		
Materiál	Relatívna permitivita $\epsilon_r$	Relatívna permeabilita $\mu_r$
sklo	5,5	1
plexisklo	3,4	1
mramor	8,3	1
tvrdá guma	3	1
ferit	12	1000



Obr. 2 Účinnosť tienenia elektromagnetického poľa pre skúmané materiály

#### IV. ZÁVER

Tento príspevok sa venuje skúmaniu účinnosti tienenia elektromagnetického poľa. V programe Ansys bol vytvorený model vlnovodu a bariéry. Bariéry boli priradené materiálové vlastnosti na základe Tab.I. V programe bolo nastavený frekvenčný rozsah od 2 GHz do 8 GHz. Následne boli určené krivky účinnosti tienenia elektromagnetického poľa pre 5 rôznych materiálov. Z výsledkov vyplýva, že zo spomínaných materiálov (sklo, plexisklo, mramor, tvrdá guma a ferit s hrúbkami 1 cm) dosahuje najvyššie hodnoty účinnosti tienenia ferit. Tento výsledok bolo možné predpokladať a teda potvrdil sa. Modelovaním v programe Ansys boli overené tieto predpoklady a správnosť navrhnutého modelu vlnovodu a bariéry.

#### POĎAKOVANIE

Tento príspevok vznikol vďaka podpore projektu VEGA No. 1/0435/19.

#### REFERENCES

- [1] J. Durpok, H-J. MORISKE, and N. Englert, "GESÜNDER WOHNEN — ABER WIE?: Praktische Tipps für den Alltag," Berlin: KOMAG mbH, 2005. [online]. [citované 2015-07-07]. Dostupný na internete: < <http://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/2885.pdf>>.
- [2] ICNIRP STATEMENT, "On the 'Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz)'" [online]. [Dátum: 18. 09. 2016.] Dostupné na internete: <http://www.icnirp.org/cms/upload/publications/ICNIRPStatementEMF.pdf>
- [3] Ha, M. - Lim, HJ. - Cho, SH. - Choi, HD. - Cho, KY.: "Incidence of cancer in the vicinity of Korean AM radio transmitters" [online]. [Dátum: 23. 03. 2016.] Dostupné na internete: <http://www.ncbi.nlm.nih.gov/pubmed/15859510>
- [4] 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
- [5] IEEE Standard 299-2006, Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures, EMC Society, New York 2006, p.39

# Návrh inteligentného požiarneho varovného systému s využitím neurónových fuzzy prostriedkov

Ján VAŠČÁK

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

jan.vascak@tuke.sk

**Abstract** – Problematika falošných požiarnych hlásení predstavuje okrem iného aj vážny ekonomický problém. Preto je potrebné aj vyhodnotiť stupeň rizika požiaru a sledovať v čase. Predmetom tohto článku je popis návrh lacného a samostatného systému varovania pred požiarom s využitím prostriedkov ako Arduino a prostriedkov umelej inteligencie na vyhodnocovanie rizika. V tomto prípade sa použili modely založené na sieťach ANFIS a RBF. V práci je uvedený popis návrhu, tréningu modelov a experimentov vykonaných na nich, spolu s ich vyhodnotením.

**Keywords** – ANFIS, falošný poplach, požiarne hlásič, RBF

## I. ÚVOD

Na trhu sa sice nachádzajú hlásiče požiaru a výstražné požiarne systémy založené na kombinácii dymového senzora a alarmu, avšak takéto systémy požiarnej signalizácie sú niekedy náchylné na chyby a môžu reagovať na neaktuálne ukazovatele prítomnosti požiaru, klasifikované ako falošné varovania. Tieto môžu spôsobovať najmä vo veľkých mestách vysoké finančné straty, niekde až do výšky desiatok miliónov Eur ročne (napr. v Londýne je to vyše 40 miliónov Eur) [1]. Falošné alarmy sa môžu spustiť z rôzneho dôvodu, ktorý závisí od typu detektora. Najčastejšie je to pod vplyvom environmentálneho faktora. Z tohto dôvodu je veľmi užitočné vyvíjať nízkorozpočtový detektor požiaru, ktorý tieto falošné požiare redukuje [2].

## II. POPIS PROBLEMATIKY A PROSTRIEDKOV

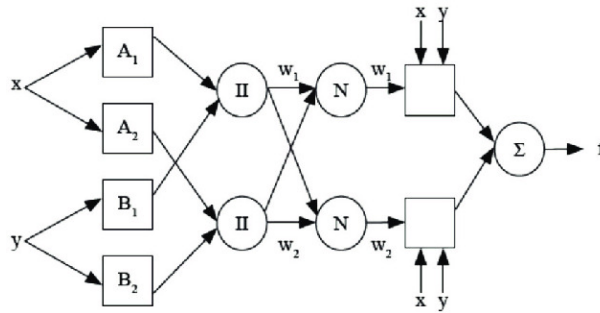
Pre potreby návrhu inteligentného zariadenia na detekciu požiaru s minimom falošných poplachov, ktoré by bolo aj cenovo dostupné, sa využili hardvérové prostriedky pozostávajúce z dosky Arduino Uno a týchto štyroch snímačov:

- snímač dymu,
- snímač plameňov,
- snímač teploty,
- snímač vlhkosti.

Za pomoci týchto snímačov bol vytvorený dátový súbor pre učenie dvoch modelov, ktoré budú predikovať pravdepodobnosť toho, že vznikne požiar. Ako referenčný model bol zvolený adaptívny neurónový fuzzy inferenčný systém ANFIS a pre účely porovnania neurónové siete založené na radiálnych funkciách, tzv. RBF siete. V ďalšom budú tieto prostriedky stručne popísané.

### A. ANFIS

Adaptívny neurónový fuzzy inferenčný systém ANFIS je druh umelej neurónovej siete využívajúci princíp inferencie tzv. TSK (Takagi-Sugeno-Kang) fuzzy regulátora. Pretože integruje neurónové siete aj princípy fuzzy logiky, má potenciál zachytiť výhody oboch typov prostriedkov v jednom rámci. Jeho inferenčný systém zodpovedá množine fuzzy pravidiel AK – POTOM, ktoré majú schopnosť učiť sa aproximovať nelineárne funkcie [3]. Jeho všeobecná architektúra je uvedená na Obr. 1.



Obr. 1 Architektúra modelu ANFIS.

**B. RBF**

Radiálna bázová funkcia, skrátene radiálna funkcia, (RBF) je funkcia so skutočnou hodnotou  $\varphi$ , ktorej hodnota závisí iba od vzdialenosti medzi vstupom a niektorým fixným bodom  $c$  (napr. počiatkom súradnicového systému), čiže  $\varphi(x) = \varphi(\|x\|)$ , resp.  $\varphi(x) = \varphi(\|x - c\|)$ . Akákoľvek funkcia  $\varphi$ , ktorá vyhovuje vlastnosti  $\varphi(x) = \varphi(\|x\|)$ , je radiálna funkcia. Vzdialenosť je zvyčajne euklidovská, aj keď niekedy sa používajú aj iné metriky (napr. manhattanská). Často sa používa množina takýchto funkcií  $\{\varphi_k\}_k$ , ktorá vytvára základ pre návrh tzv. RBF neurónových sietí [4], avšak tento typ funkcií je možné použiť napr. aj v metóde tzv. podorných vektorov.

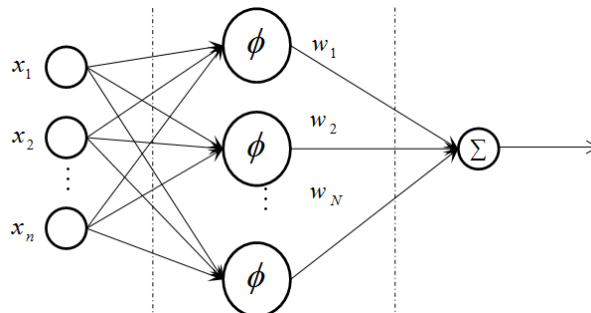
Výstup RBF siete z Obr. 2 je možné definovať ako vážený súčet výstupov zo skrytej vrstvy:

$$y(x) = \sum_{i=1}^N w_i \varphi(|x - x_i|), \tag{1}$$

ktorá sa nazýva aj receptívne pole a obsahuje v sebe sadu radiálnych funkcií preberajúcich úlohu aktivačných funkcií jedinej skrytej vrstvy siete. Je možné preukázať, že ktorúkoľvek spojitú funkciu na kompaktnom intervale, je možné v zásade interpolovať s ľubovoľnou presnosťou súčtom tohto tvaru, ak sa použije dostatočne veľký počet radiálnych funkcií  $N$ .

Keďže výstup zo siete  $y(x)$  je diferencovateľný vzhľadom na váhy  $w_i$ , tak sa tieto váhy dajú naučiť pomocou ktorejkoľvek zo štandardných iteračných metód pre neurónové siete. V tomto prípade boli váhy medzi skrytou vrstvou a výstupnou vrstvou vypočítané pomocou Moore-Penrosovej inverzie [5]. Tento algoritmus prekonáva mnoho problémov v tradičných gradientových algoritmoch, ako je kritérium zastavenia, rýchlosť učenia, počet epoch a miestne minimá. Vďaka kratšiemu tréningovému času a schopnosti zovšeobecňovania je vhodný pre aplikácie v reálnom čase.

Ako radiálna funkcia môže byť použitá ľubovoľná, ktorá je stredovo súmerná, ale v ďalšom sa bude využívať gaussovská funkcia, nakoľko stred a distribúcia aktivačných funkcií by mali mať spravidla podobné charakteristiky ako štatistické charakteristiky tréningových dát a vo väčšine prípadov sa jedná práve o gaussovské rozdelenie. Na základe univerzálnej teóremy aproximácie a distribúcie aktivačných funkcií toto nie je deterministické, ak je počet skrytých neurónov dostatočný, čiže sieť spätnej väzby s jednou skrytou vrstvou s dostatočným počtom skrytých neurónov dokáže aproximovať ľubovoľnú funkciu na ľubovoľnú úroveň presnosti.



Obr. 2 Architektúra RBF siete.

### C. Dátový súbor

Pre potreby tréningovania použitých neurónových sietí bol vytvorený vlastný dátový súbor, ktorý obsahuje 5 veličín:

- čas [min.],
- zmena teploty [ $^{\circ}\text{C}$ ],
- zmena vlhkosti [%],
- zmena zadymenia [ppm],
- výstup zo senzora plameňov (0–nehorí, 1–horí).

Dáta boli získavané v reálnom čase pomocou jednodoskového regulátora Arduino a nástroja PLX-DAQ v programe Microsoft Excel. Následne tieto boli transformované do predpísaného formátu a zapisované do .csv súboru. Fragment takéhoto dátového súboru je Tabuľka 1.

Tabuľka 1  
 Ukážka tréningového dátového súboru.

Čas [min.]	Zmena teploty [ $^{\circ}\text{C}$ ]	Zmena vlhkosti [%]	Zmena zadymenia [ppm]	Senzor plameňa [0/1]	Riziko požiaru
1.26	0.5	0.6	5	0	nízke
1.35	0.5	1.4	2	0	nízke
1.80	0.7	4.1	1	0	stredné
2.41	3.2	7.4	2	0	stredné
1.28	1.7	0.9	10	0	stredné
2.28	4.7	0.7	3	0	stredné
1.10	3.7	3.4	3	0	stredné
0.60	0.2	1.1	5	0	nízke
1.10	0.7	0.5	62	1	vysoké
...	...	...	...	...	...

### III. NÁVRH A REALIZÁCIA

Pre implementáciu oboch vyššie spomenutých modelov bol využitý programovací prostriedok Arduino Uno s príslušnými knižnicami a pre vizualizáciu výsledkov bola vytvorená jednoduchá obrazovka v programe C#.

V prípade modelu využívajúceho ANFIS je vytvorených 81 pravidiel so štyrmi vstupnými veličinami (teplota, vlhkosť, dym a čas) po troch lingvistických hodnotách, z ktorých je vygenerovaný výstup ako miera rizika požiaru v rozmedzí [0; 1]. Časové intervaly boli odhadované empiricky a sú zobrazené v Tabuľke 2 a rozdelenie úrovní rizika požiaru je zobrazené v Tabuľke 3.

Tabuľka 2  
 Definície časových intervalov.

	Časový interval [min.]
veľmi krátky	0 - 0.5
krátky	0.5 - 1.5
dlhý	1.5 - 2.5

Tvorba modelu ANFIS bola započatá generovaním pravidiel ako súbor kombinácií vstupných lingvistických hodnôt. Následne bol model tréningovaný s využitím dátového súboru, ktorý celkovo obsahoval 75 záznamov. Maximálny počet epoch pri tréningovaní bol nastavený na 60, avšak sa ukázalo, že navrhnutý model dosiahol najnižšiu chybovú funkciu už pri hranici okolo 35 epoch. Po úspešnom natréningovaní modelu navrhnutý model používal na evaluáciu dáta prichádzajúce z Arduina. Výstupom evaluácie je rozsah hodnôt 0–1, ktoré sa zapisujú do súboru.

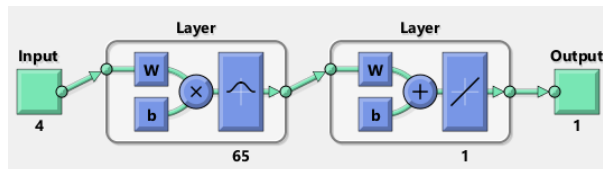
Tabuľka 3  
 Definície úrovní rizika požiaru.

Úroveň rizika	Teplota [ $^{\circ}\text{C}$ ]	Vlhkosť [%]	Zadymenie [ppm]
nízka	0 - 1.5	0 - 3.0	0 - 8
stredná	1.0 - 5.0	2.0 - 8.0	6 - 50
vysoká	4.0 - 10.0	6.0 - 20.0	40 - 100

Podobne sa postupovalo pri vytvorení RBF modelu, v rámci ktorého sa optimalizujú nasledujúce parametre:

- stredná kvadratická odchýlka chybovej funkcie,
- rozpätie radiálnych funkcií,
- maximálny počet epoch.

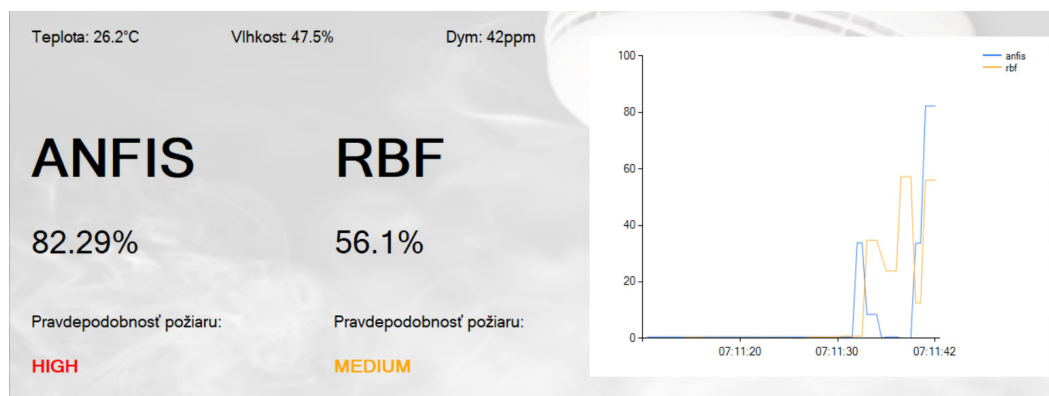
Aj v tomto prípade bolo cieľom dosiahnuť čo najnižšiu chybovú funkciu a preto bol prvý parameter nastavený na 0. Druhý parameter určuje plynulosť aproximácie funkcie. Príliš veľké rozpätie znamená, že na prispôbenie rýchlo sa meniacej funkcii je potrebných veľa neurónov. Príliš malá šírka znamená, že na zabezpečenie hladkej funkcie je potrebných veľa neurónov a sieť sa nemusí dobre zovšeobecňovať. Tento parameter bol pôvodne prednastavený na hodnotu 1. Avšak také nastavenie nevyhovovalo, pretože skok medzi jednotlivými meraniami bol príliš veľký. Po viacerých pokusoch boli najlepšie výsledky dosahované pri hodnote 1,865. Podobne ako pri rozptyle, aj v prípade počtu epoch bolo pokusmi zistené, že maximálny počet epoch by nemal prekračovať hodnotu 65. Na Obr. 3 je znázornený výsledný RBF model.



Obr. 3 Architektúra RBF modelu.

#### IV. VYHODNOTENIE EXPERIMENTOV

Výsledky navrhnutých modelov boli testované a porovnávané pomocou evaluačného programu vytvoreného v prostredí Matlabu a pomocou vytvorenej C# aplikácie. Program v Matlabe slúžil na komunikáciu s Arduinom cez sériový port. Cez tento port boli v Matlabe spracovávané všetky merané veličiny zo senzorov, ktoré boli na Arduine pripojené. Zároveň sa v cykle overovalo, či nedochádza k zmene aspoň jednej z meraných veličín. Pokiaľ áno, začal sa počítat interval v minútach a každú sekundu sa pomocou navrhnutých modelov vyhodnocovala pravdepodobnosť požiaru, pričom vstupom boli zmeny veličín za určitý čas. Vyhodnocovanie prebiehalo iba pokiaľ boli zmeny veličín zaznamenávané kratšie ako 10 sekúnd. Výsledky týchto vyhodnotení a výstupy zo senzorov boli ukladané do textového súboru, z ktorého čítala dáta C# aplikácia. Jej hlavnou úlohou bola vizualizácia výstupov. Táto aplikácia zobrazovala výstupy zo senzorov a výsledky vyhodnotení získané natrénovanými modelmi ANFIS a RBF. Aplikácia obsahovala aj trend, ktorý zobrazoval vypočítané pravdepodobnosti požiaru za poslednú minútu pre obidva modely. Poslednou časťou, ktorú táto vizualizácia obsahuje, je vyhodnotenie rizika požiaru, kde označenie *nízke* zodpovedá percentuálnej pravdepodobnosti [0; 20], *stredné* [21; 50] a *vysoké* [51; 100]. Výsledná aplikácia vizualizácie výstupov z obidvoch modelov je na Obr. 4.

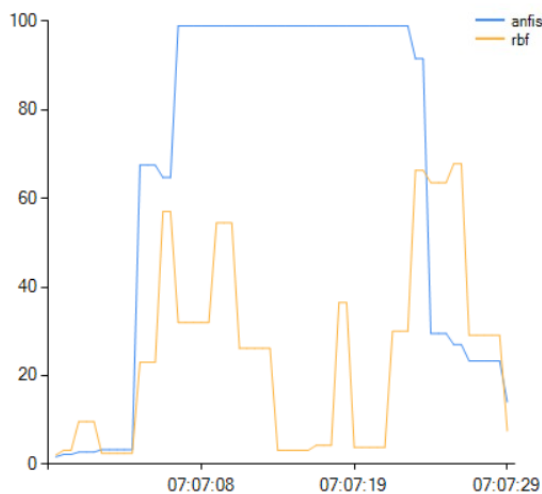


Obr. 4 Ukážka vizualizácie požiarneho varovného systému.

Experimenty prebiehali pomocou rôznych spôsobov navodzujúcich požiar ako napríklad pomocou zápalek, zapaľovača či tvorenia dymu. Jednalo sa o podobné metódy, akými sa vytváral aj tréningový dátový súbor. Vyhodnocovanie sa sústreďovalo hlavne na to, ako kvalitne a rýchlo dokážu jednotlivé modely vyhodnotiť požiar v prípade, ak sa v prostredí požiar naozaj vyskytuje a následne sa pozornosť zamerala na odolnosť voči falošným požiarom.

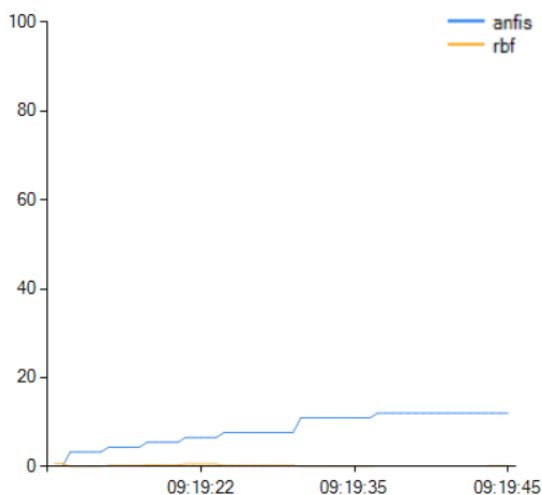
V počiatočných experimentoch sa vyhodnocovala reakcia na skutočný požiar, resp. reakcia na prítomnosť ohňa v blízkosti senzorov. Z väčšiny výsledkov bolo zrejmé, že ANFIS model bol rýchlejší v detekcii požiaru, a to v priemere o 2 – 3 sekundy. Z priebehu na Obr. 5 je takisto

zrejme, že model ANFIS bol aj oveľa citlivejši ako RBF, a to do takej miery, že RBF model iba zriedka vyhodnotí požiar na najvyššiu pravdepodobnosť. Podobné výsledky dosahujú oba modely iba, ak sa jedná o postupné zvyšovanie teploty, či dymu. V takom prípade sú výsledky približne na rovnakej úrovni, avšak RBF takmer vždy udáva menšiu pravdepodobnosť ako ANFIS.



Obr. 5 Experiment pri požari.

V ďalších sa pozornosť zamerala na odolnosť voči falošnému požiaru, a to takým spôsobom, že senzory boli vystavené prostriedkom, ktoré generujú teplo, alebo dym, no iba v malých dávkach a intervaloch ako napr. šálka kávy, či vytváranie malého množstva dymu. Experimenty ukázali, vid' Obr. 6, že pokiaľ sa jednálo o zvýšenie teploty o  $0.5\text{ }^{\circ}\text{C}$ , zmenu vlhkosti 1,5% a nulovej zmeny ppm, tak pri intervale 2 minút pravdepodobnosť požiaru dosiahne pri ANFIS modeli najviac 12%, zatiaľčo pri RBF modeli iba 3%. Pri zmene teploty o  $2\text{ }^{\circ}\text{C}$  a vlhkosti 2.7% je však za 3 minúty pravdepodobnosť už 38% pri ANFIS modeli a 6% pri RBF modeli. Pokiaľ sa jednálo o krátkodobú vysokú zmenu teploty, tak pravdepodobnosť požiaru bola príliš vysoká, a tak by v tomto prípade boli výsledky zo siete ANFIS relevantné iba po viac ako 5 sekundách od detekcie vysokej pravdepodobnosti požiaru, keďže výsledky po určitom čase naspäť klesnú na nízke hodnoty.



Obr. 6 Experiment pri falošnom požari.

## V. ZÁVER

V rámci tohto článku bol prezentovaný vlastný dátový súbor, ktorý bol vytvorený pomocou senzorov teploty, vlhkosti, dymu a plameňa pri rôznych situáciách, ktoré indikovali existenciu

alebo neexistenciu požiaru. Tento dátový súbor bol použitý na tréovanie modelov ANFIS a RBF, ktoré boli implementované na Arduine, a tým bola zabezpečená samostatná funkčnosť takéhoto zariadenia bez nutnosti pripojenia na ďalšie zariadenie, či počítač. Po definovaní najlepších parametrov navrhnutých modelov, tieto slúžili na vyhodnocovanie požiaru v reálnom čase. Vyhodnocovanie úspešnosti prebiehalo pomocou evaluačného programu prostredia Matlab a C# aplikácie na vizualizáciu výstupov, teda najmä pravdepodobnosti výskytov požiarov jednotlivých modelov. Z výsledkov vyplýva, že model ANFIS bol vhodnejší pre aplikáciu inteligentného detektora požiaru oproti modelu RBF, a to najmä kvôli vyššej rýchlosti detekcie a vyššej citlivosti. Oba modely sú však vhodné na detekciu postupne sa rozrastajúceho požiaru a sú aj sčasti odolné voči falošným požiarom. V ďalšom pre zlepšenie výsledkov by bolo vhodné jednak zväčšiť tréovací dátový súbor pre ďalšie možné situácie a jednak pridať posielanie upozornení pre užívateľa v prípade detekcie požiaru, ako sa to robí napr. v systémoch starostlivosti zdravia [6].

#### LITERATÚRA

- [1] B. Sarwar, I. S. Bajwa, N. Jamil, S. Ramzan, and N. Sarwar, "An intelligent fire warning application using IoT and an adaptive neuro-fuzzy inference system," *Sensors*, vol. 19, no. 14, 2019.
- [2] M. A. Jallal, A. González-Vidal, A. F. Skarmeta, S. Chabaa, and A. Zeroual, "A hybrid neuro-fuzzy inference system-based algorithm for time series forecasting applied to energy consumption prediction," *Applied Energy*, vol. 268, pp. 1–19, 2020.
- [3] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [4] Y. Wu, H. Wang, B. Zhang, and K.-L. Du, "Using radial basis function networks for function approximation and classification," *International Scholarly Research Notices*, pp. 1–34, 2012.
- [5] A. Ben-Israel and T. N. Greville, *Generalized Inverses*, 2nd ed., ser. CMS Books in Mathematics. Springer, New York, NY, 2003.
- [6] K. Hameed, I. S. Bajwa, S. Ramzan, W. Anwar, and A. Khan, "An intelligent iot based healthcare system using fuzzy neural networks," *Scientific Programming*, pp. 1–15, 2020.



# Riadenie BLDC motora

<sup>1</sup>Peter GIROVSKÝ, <sup>2</sup>Martin ZÁVIŠ

<sup>1,2</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Letná 1/9, 04200 Košice-Sever, Slovenská republika

<sup>1</sup>peter.girovsky@tuke.sk, <sup>2</sup>martin.zavis@student.tuke.sk

**Abstrakt** — V tomto príspevku je prezentovaný menič a riadenia pre BLDC motor. V článku je prezentovaný simulačný model BLDC motora a meniča pre jeho napájanie. V závere príspevku sú prezentované dosiahnuté simulačné výsledky riadenia BLDC motora. Celý návrh riadenia BLDC motora a jeho následné simulačné overenie bolo robené v programe MATLAB.

**Kľúčové slová** — BLDC motor, simulácia, riadenie, menič

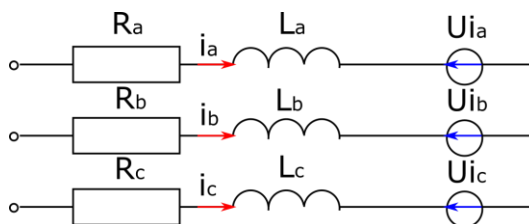
## I. ÚVOD

V súčasnosti sú elektrické motory všade tam, kde je potrebné premeniť elektrickú energiu na mechanickú energiu ako napríklad vo výrobe, v robotike a elektrických autách. Vývoj a výskum v oblasti elektrických motorov sa uberať skôr zdokonaľovaním súčasných typov riadenia, zefektívňovaním účinnosti motorov i meničov a ich a miniaturizáciou. BLDC motory v súčasnosti začínajú pomaly vytláčať klasické jednosmerné motory s komutátorom pre výborné vlastnosti BLDC motora ako napríklad ich vysokú účinnosť, spoľahlivosť a bezúdržbovosť.

Cieľom tohto príspevku je predstaviť simulačný model BLDC motora, meniča pre jeho napájanie a spôsob riadenia BLDC motora. Tento simulačný návrh bol následne otestovaný, pričom výsledky simulácií v programe MATLAB sú prezentované v závere tohto článku.

## II. MATEMATICKÝ MODEL BLDC MOTORA

Matematický model BLDC motora vychádza zo schémy zobrazenej na obr.1. Jednotlivé fázy motora sú označené indexami a, b, c. R predstavuje odpor jednej fázy vinutia motora, L je indukčnosť a  $U_i$  predstavuje indukované napätie.



Obr. 1 Schéma statorového vinutia BLDC motora.

Na základe tejto schémy BLDC motora môžeme napísať matematické rovnice troj-fázového BLDC motora.

$$u_{ab} = R(i_a - i_b) + L \frac{d}{dt}(i_a - i_b) + (U_{i_a} - U_{i_b}) \quad (1)$$

$$u_{bc} = R(i_b - i_c) + L \frac{d}{dt}(i_b - i_c) + (U_{i_b} - U_{i_c}) \quad (2)$$

$$u_{ca} = R(i_c - i_a) + L \frac{d}{dt}(i_c - i_a) + (U_{i_c} - U_{i_a}) \quad (3)$$

Pre jednotlivé indukované napätia platia nasledujúce rovnice:

$$U_{i_a} = \frac{k_{U_i}}{2} \omega_m F(\theta_e) \quad (4)$$

$$U_{i_b} = \frac{k_{U_i}}{2} \omega_m F\left(\theta_e - \frac{2\pi}{3}\right) \quad (5)$$

$$U_{i_c} = \frac{k_{U_i}}{2} \omega_m F\left(\theta_e - \frac{4\pi}{3}\right) \quad (6)$$

Kde  $k_{U_i}$  predstavuje konštantu indukovaného napätia a  $\Theta_e$  elektrický uhol rotora. Momentová rovnica môžeme napísať v tvare:

$$M_{em} - k_t \omega_m - M_z = J \frac{d\omega_m}{dt} \quad (7)$$

V tejto rovnici hodnota  $k_t$  predstavuje konštantu trenia  $M_z$  predstavuje záťažný moment a pre elektromagnetický moment motora  $M_{em}$  platí vzťah:

$$M_{em} = \frac{k_m}{2} \left[ F(\theta_e) i_a + F\left(\theta_e - \frac{2\pi}{3}\right) i_b + F\left(\theta_e - \frac{4\pi}{3}\right) i_c \right] \quad (8)$$

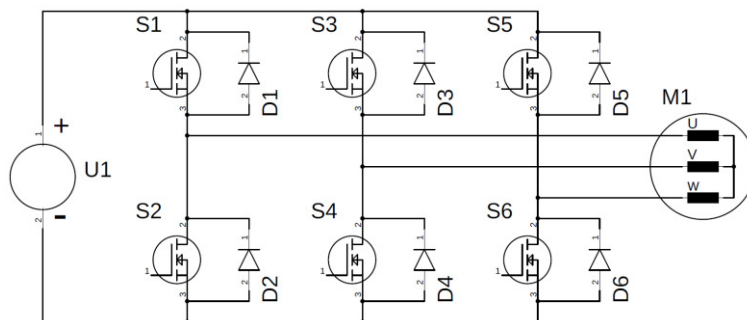
Pre zjednodušenie modelu BLDC motora upravíme rovnice 1-3 dostávame rovnice BLDC motora, na základe ktorých bol zostavený jeho simulačný model.

$$u_{ab} = R(i_a - i_b) + L \frac{d}{dt}(i_a - i_b) + (U i_a - U i_b) \quad (9)$$

$$u_{bc} = R(i_a + 2i_b) + L \frac{d}{dt}(i_a + 2i_b) + (U i_b - U i_c) \quad (10)$$

### III. MATEMATICKÝ MODEL MENIČA PRE NAPÁJANIE BLDC MOTORA

Pre napájanie BLDC motor bol použitý 3-fázový mostík zobrazený na obr. 2. Vstupná hodnota napätia motora je vypočítavaná na základe aktuálneho elektrického uhla rotora a veľkosti zodpovedajúceho prúdu podľa tabuľky 1.



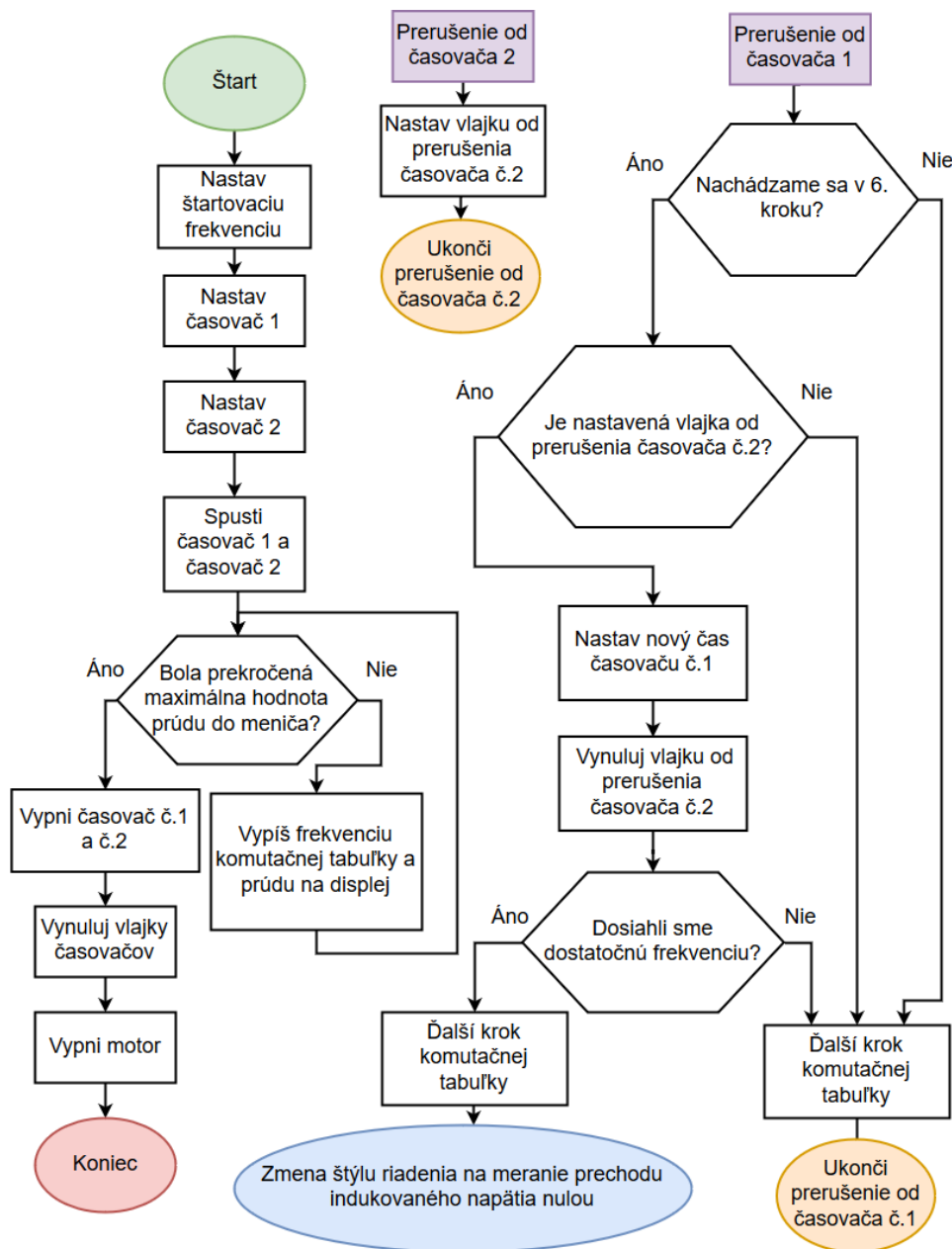
Obr. 2 Schéma 3-f mostíka pre napájanie BLDC motora.

Table 1  
 Rovnice pre výpočet vstupného napätia LDC motora

Elektrický uhol	Prúd diódy	$u_{ab}-U_{i_{ab}}$	$U_{bc}-U_{i_{bc}}$
$0^\circ-60^\circ$	$i_a \neq 0, i_c \neq 0$	$-U_{i_a} + U_{i_b}$	$-U_{i_b} + U_{i_c}$
	$i_a \neq 0, i_c = 0$	$-U_{i_a} + U_{i_b}$	$\frac{1}{2}(U_{i_a} - U_{i_b})$
	$i_a = 0, i_c \neq 0$	$\frac{1}{2}(U_{i_b} - U_{i_c})$	$-U_{i_b} + U_{i_c}$
	$i_a = 0, i_c = 0$	0	0
$60^\circ-120^\circ$	$i_a \neq 0, i_b \neq 0$	$-U_{i_a} - U_{i_b} + U_{i_c}$	$U_{i_b} - U_{i_c}$
	$i_a \neq 0, i_b = 0$	$\frac{1}{2}(-U_{i_a} + U_{i_c})$	$\frac{1}{2}(-U_{i_a} + U_{i_c})$
	$i_a = 0, i_b \neq 0$	$\frac{1}{2}(-U_{i_a} - U_{i_b} - U_{i_c})$	$U_{i_b} - U_{i_c}$
	$i_a = 0, i_b = 0$	0	0
$120^\circ-180^\circ$	$i_b \neq 0, i_a \neq 0$	$-U_{i_a} + U_{i_b}$	$-U_{i_b} + U_{i_c}$
	$i_b \neq 0, i_a = 0$	$\frac{1}{2}(U_{i_b} - U_{i_c})$	$-U_{i_b} + U_{i_c}$
	$i_b = 0, i_a \neq 0$	$\frac{1}{2}(-U_{i_a} + U_{i_c})$	$\frac{1}{2}(-U_{i_a} + U_{i_c})$
	$i_b = 0, i_a = 0$	0	0
$180^\circ-240^\circ$	$i_b \neq 0, i_c \neq 0$	$-U_{i_a} + U_{i_b}$	$-U_{i_b} - U_{i_c}$
	$i_b \neq 0, i_c = 0$	$-U_{i_a} + U_{i_b}$	$\frac{1}{2}(U_{i_a} - U_{i_b})$
	$i_b = 0, i_c \neq 0$	$\frac{1}{2}(-U_{i_a} - U_{i_b} + U_{i_c})$	$\frac{1}{2}(-U_{i_a} - U_{i_b} + U_{i_c})$
	$i_b = 0, i_c = 0$	0	0
$240^\circ-300^\circ$	$i_c \neq 0, i_b \neq 0$	$-U_{i_a} + U_{i_b}$	$-U_{i_b} + U_{i_c}$
	$i_c \neq 0, i_b = 0$	$\frac{1}{2}(-U_{i_a} + U_{i_c})$	$\frac{1}{2}(-U_{i_a} + U_{i_c})$
	$i_c = 0, i_b \neq 0$	$-U_{i_a} + U_{i_b}$	$\frac{1}{2}(U_{i_a} - U_{i_b})$
	$i_c = 0, i_b = 0$	0	0
$300^\circ-360^\circ$	$i_c \neq 0, i_a \neq 0$	$U_{i_a} - U_{i_b} + U_{i_c}$	$-U_{i_b} + U_{i_c}$
	$i_c \neq 0, i_a = 0$	$\frac{1}{2}(U_{i_b} - U_{i_c})$	$-U_{i_b} + U_{i_c}$
	$i_c = 0, i_a \neq 0$	$U_{i_a} - U_{i_b} + U_{i_c}$	$\frac{1}{2}(-U_{i_a} - U_{i_b})$
	$i_c = 0, i_a = 0$	0	0

#### IV. NÁVRH RIADENIA BLDC MOTORA

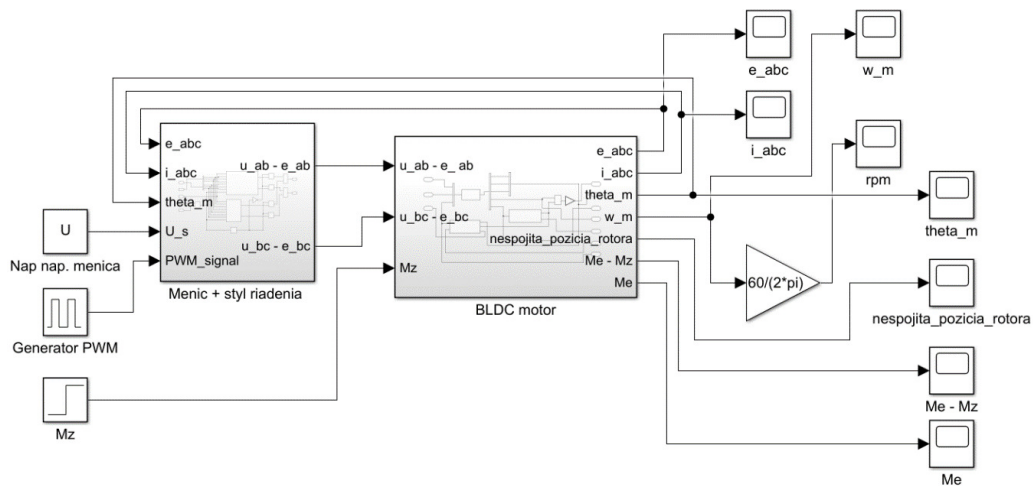
V našom prípade pre sme per určenie polohy rotora použili metódu, pri ktorej sa sníma indukované napätie. Keďže napätie sa indukuje pri otáčaní rotora a závisí od jeho uhlovej rýchlosti, tak pri štarte motora, keď sa rotor netočí nevieme zistiť polohu rotora pretože indukované napätie je nulové. Motor preto budeme preto spúšťať v otvorenej slučke a fázy motora spínané na základe časových intervalov. Po nábehu motora sa spínanie tranzistorov bude riadiť na základe indukovaného napätia. Principiálny algoritmus navrhnutého riadenia BLDC motora je zobrazený na obr. 3.



Obr. 3 Algoritmus riadenia BLDC motora.

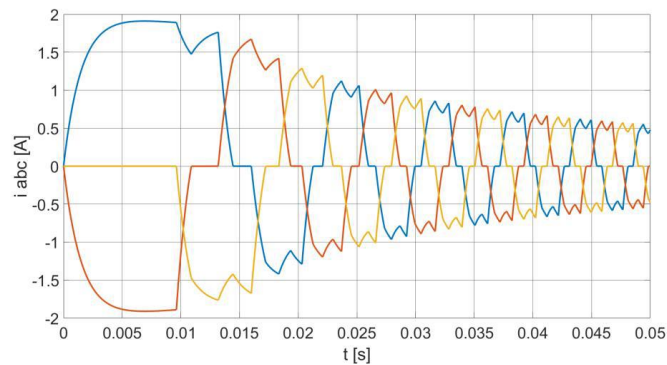
#### V. VERIFIKÁCIA SIMULAČNÉHO MODELU BLDC MOTORA

Na základe matematických rovníc BLDC motora bol zostavený jeho simulačný model v programe MATLAB/Simulink, zobrazený na obr. 4. Pre simuláciu bol použitý BLDC motor, ktorého s parametrami:  $P_N=150\text{W}$ ;  $U_N=150\text{V}$ ;  $n_N=3000\text{ot/min}$ ;  $M_N=0,47\text{Nm}$ ;  $J=1,5 \cdot 10^{-3}\text{kg/m}^2$ ;  $p=2$ ;  $R=37,77\Omega$ ;  $L=5,35\text{mH}$ ;  $k_m=0,47\text{Nm/A}$ ;  $k_t=1,725 \cdot 10^{-6}\text{N}$ ;  $k_{U_i}=0,06812\text{Vs/rad}$ .



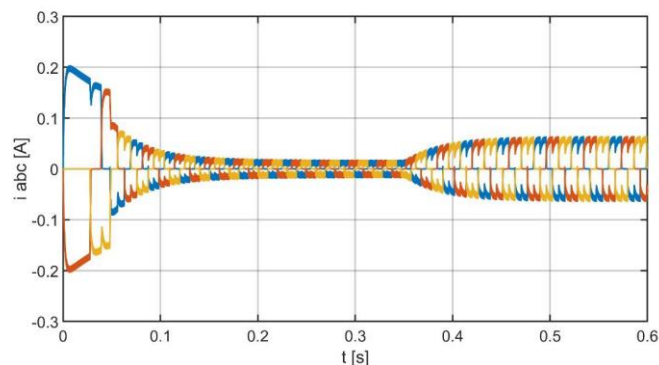
Obr. 4 Simulačný model BLDC motora s napájaním a riadením v programe MATLAB/Simulink.

Výsledky z verifikácie simulačného modelu BLDC motora sú zobrazené na nasledujúcich obrázkoch. Na obr. 5 je znázornený priebeh prúdov pri nábehu BLDC motora, pri napájaní nominálnym napätím 150V.

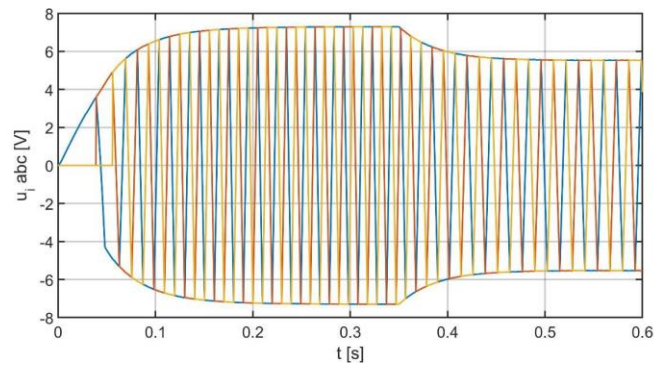


Obr. 5 Priebehy prúdu pri rozbehu BLDC motora bez záťaže.

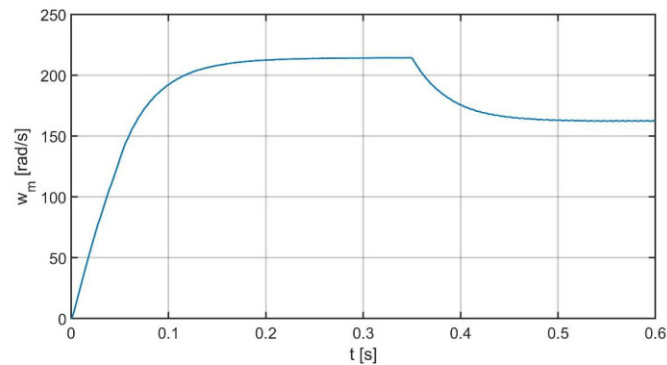
Na nasledujúcich priebehoch sú znázornené priebehy zo simulácie BLDC motora pri napájaní zníženým napätím, približne 10% hodnoty nominálneho napätia.



Obr. 6 Priebehy prúdu pri rozbehu BLDC motora pri zníženom napätí, v čase 0,35s bol motor zaťažný momentom 0,02Nm.



Obr. 7 Priebehy indukovaného napätia BLDC motora pri zníženom napätí, v čase 0,35s bol motor zaťažovaný momentom 0,02Nm.



Obr. 8 Priebeh uhlovej rýchlosti BLDC motora pri zníženom napätí, v čase 0,35s bol motor zaťažovaný momentom 0,02Nm.

## VI. ZÁVER

V tomto článku bol prezentovaný simulačný model BLDC motora spolu s meničom pre jeho napájanie a s riadením. Výsledky simulácie uvedené v závere článku prezentujú správnosť navrhnutého simulačného modelu. Tento simulačný model BLDC motora je možné využiť pri ďalšom výskume nových progresívnych metód riadenia BLDC motora.

## POĎAKOVANIE

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-19-0210

## LITERATÚRA

- [1] R. Krisham, "Permanent Magnet Synchronous and Brushless DC Motor Drives" CRC Press Taylor & Francis group, 2010, ISBN 978-0-8247-5384-9.
- [2] D. Hanselman, „Brushless Motors“, E-Man Press LLC, 2012, ISBN 978-0-9826-9261-5
- [3] S. Baldursson. Brushless DC Motor Fundamentals Application note. [Online] Dostupné na internete: <http://webfiles.portal.chalmers.se/et/MSc/BaldurssonStefanMSc.pdf>.
- [4] I. O. Bucak. Brushless Position Error Compensation via a Variable Reluctance Sensor Applied to a Hybrid Vehicle Electric Machine. [Online] Dostupné na internete: <[https://www.researchgate.net/figure/Variable-Reluctance-VR-sensor-that-senses-movement-of-the-toothed-wheel-past-point-of\\_fig1\\_42539593](https://www.researchgate.net/figure/Variable-Reluctance-VR-sensor-that-senses-movement-of-the-toothed-wheel-past-point-of_fig1_42539593)>.

# Návrh algoritmu pre rozpoznávanie obrazu

<sup>1</sup>Peter GIROVSKÝ, <sup>2</sup>Stanislav MAČÁK

<sup>1,2</sup> Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Letná 1/9, 04200 Košice-Sever Slovenská republika

<sup>1</sup>peter.girovsky@tuke.sk, <sup>2</sup>stanislav.macak@student.tuke.sk

**Abstrakt** — V tomto príspevku je prezentovaný návrh algoritmu pre rozpoznávanie obrazu. Návrh algoritmu pre rozpoznávanie objektov na základe farby a tvaru je postavený na umelých neuronových sieťach. V tomto prípade bola pri návrhu algoritmu rozpoznávania objektov použitá konvolučná neuronová sieť. Učenie neurónovej siete a testovanie navrhnutého algoritmu bolo robené v programe MATLAB.

**Kľúčové slová** — počítačové videnie, rozpoznávanie obrazu, neuronové siete, algoritmus, simulácia

## I. ÚVOD

Počítačové videnie je oblasť počítačovej vedy, ktorá sa zameriava na napodobňovanie ľudského videnia. Umožňuje počítačom identifikovať a spracovať objekty v obrazoch a videách rovnakým spôsobom, ako to robia ľudia. V oblasti automatizácie sa s ním môžeme stretnúť najčastejšie v kontrolných a riadiacich procesoch. Využíva sa teda najmä na kontrolu správnosti vyhotovených výrobkov a na orientovanie v priestore. Počítačové videnie do nedávna fungovalo iba v obmedzenej kapacite kvôli svojej výpočtovej náročnosti. Vďaka pokrokom vo výpočtovej výkonnosti strojov a inováciách v umelej inteligencii, sa v tejto oblasti podarilo v posledných rokoch urobiť veľký skok a dokázalo prekonať človeka pri niektorých úlohách týkajúcich sa odhaľovania a označovania objektov.

Cieľom tohto príspevku je v krátkosti predstaviť problematiku počítačového videnia a na základe jednoduchého príkladu prezentovať spôsob návrhu algoritmu pre rozpoznávanie objektov na základe ich tvaru a farby.

## II. NÁVRH ALGORITMU PRE ROZPOZNÁVANIE OBRAZU

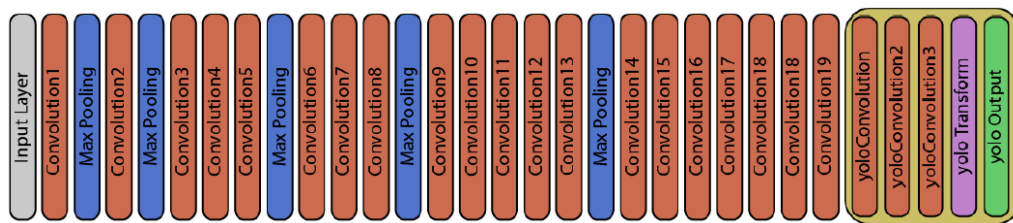
V rámci tejto práce sme sa zaoberali návrhom a testovaním algoritmu pre rozpoznávanie obrazu v reálnom čase. Návrh algoritmu pre rozpoznávanie obrazu bol založený na využití umelej neurónovej siete, pričom samotný návrh pozostával z niekoľkých krokov, ktoré sú zobrazené na obr. 1.



Obr. 1 Kroky návrhu algoritmu pre rozpoznávanie obrazu.

### A. Voľba architektúry neurónovej siete

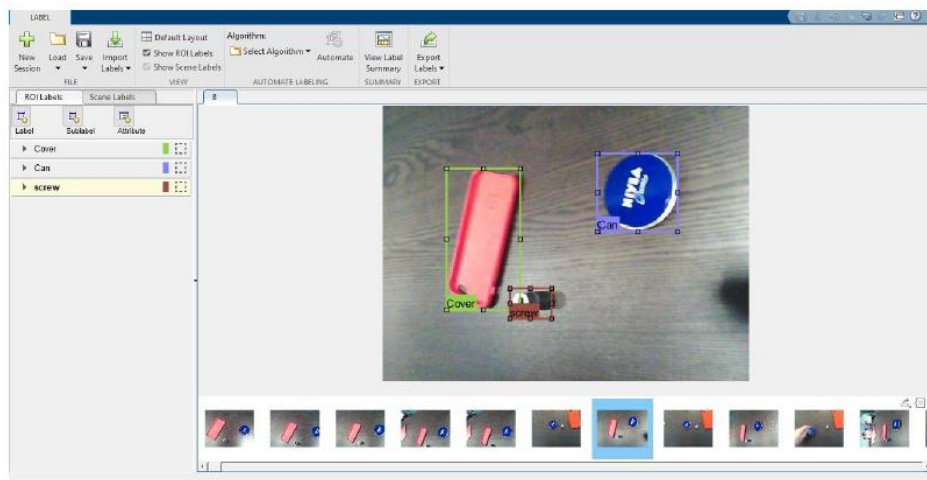
Kvôli výpočtovej nenáročnosti a dobrej presnosti bol zvolený typ neurónovej siete YOLO pre rozpoznávanie obrazu, pričom pre učenie bola použitá metóda *Transfer learning*, ktorá je založená na hĺbkovom učení. Na naučenie konkrétnej úlohy je použitý model, ktorý už bol predtým naučený na vykonávanie inej podobnej úlohy. Výhodou je, že takáto sieť je potom schopná rozoznávať príznaky v obraze s vyššími úrovňami abstrakcie.



. Obr. 2 Architektúra použitej siete.

### B. Príprava dát

Na prípravu vstupných dát bola využitá aplikácia Image Labeler, ktorá je súčasť Deep Learning toolboxu. Príprava týchto dát spočíva v nafotení predmetov, ktoré sme chceli identifikovať a ich následnom označení pomocou ohraničujúcich boxov. V procese učenia sa, porovnávajú výsledky z detekcie modelu s označenými dátami.



. Obr. 3 Príprava dát v programe Image Labeler.

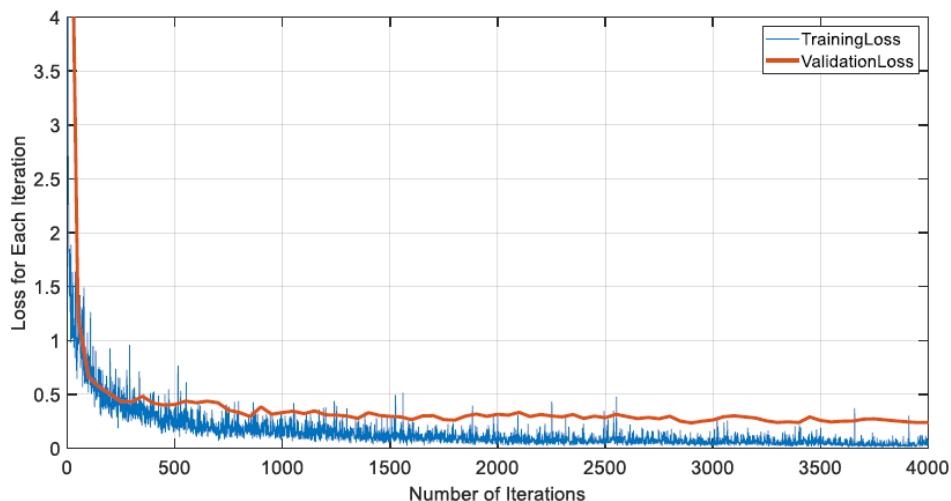
Dataset vzoriek na trénovanie modelu obsahoval 800 Obrázkov veľkosti 640x480px. Pred trénovaním boli tieto obrázky zmenšené do vhodnej veľkosti, keďže naša sieť brala na vstupe obrázky s rozlíšením 352x352px. Tento dataset sme následne rozdelili na trénovací (75%), testovací (15%) a validačný (10%) dataset.

### C. Trénovanie siete

Počas trénovania siete, model optimalizuje strednú kvadratickú chybu v chybovej funkcii medzi identifikovaným ohraničujúcim boxom a boxom označeným v GT. Chybová funkcia našej siete typu YOLO-v2 sa skladá z troch častí:

- **Lokalizačná chyba** - vyjadruje chybu od polohy a veľkosti ohraničujúceho boxu.
- **Chyba istoty** – vyjadruje chybu, či sa v danom ohraničujúcom boxe mal alebo nemal nachádzať objekt.
- **Klasifikačná chyba** – vyjadruje chybu od klasifikácie objektov.

Na optimalizovanie chybovej funkcie bola použitá metóda stochastic gradient descent (SGD). Štandardná gradientová metóda aktualizuje parametre (váhy a prahy), s cieľom minimalizovať chybovú funkciu tým, že pri každej iterácii urobí malé kroky v smere negatívneho gradientu. Gradient má smer do najväčšieho zvýšenia funkcie, a keďže musíme minimalizovať chybovú funkciu, potrebujeme ísť v smere negatívneho gradientu. Pri klasickej gradientovej metóde učenia je potrebné použiť všetky vzorky na uskutočnenie jednej iterácie učenia. Z tohto hľadiska je táto metóda učenia náročná pre výpočtový výkon. Pri stochastickej gradientovej metóde sa pri každej iterácii vypočíta gradient chybovej funkcie iba z určitej časti datasetu a následne sa aktualizujú parametre. Počas trénovania sa teda dataset rozdelí na určitý počet častí. Na prejdienie celého datasetu bude potrebný určitý počet iterácií a tie sú dané veľkosťou datasetu a veľkosťou určených častí.



. Obr. 4 Priebeh chybovej funkcie učenia

#### D. Testovanie modelu

Natrénovaná sieť bola následne testovaná pomocou testovacieho datasetu. Navrhnutý model bol trénovaný na rozpoznávanie 3 objektov, ako sú znázornené na obr. 3 a pre jednotlivé objekty bola dosiahnutá nasledujúca priemerná presnosť:

- **Objekt 1** – 94%
- **Objekt 2** – 98%
- **Objekt 3** – 84%

Presnosť jednotlivých objektov sa počítala na základe vzťahu:

$$presnosť = \frac{správne\ pozitívny}{správne\ pozitívny + falošne\ pozitívny} \quad (1)$$

kde:

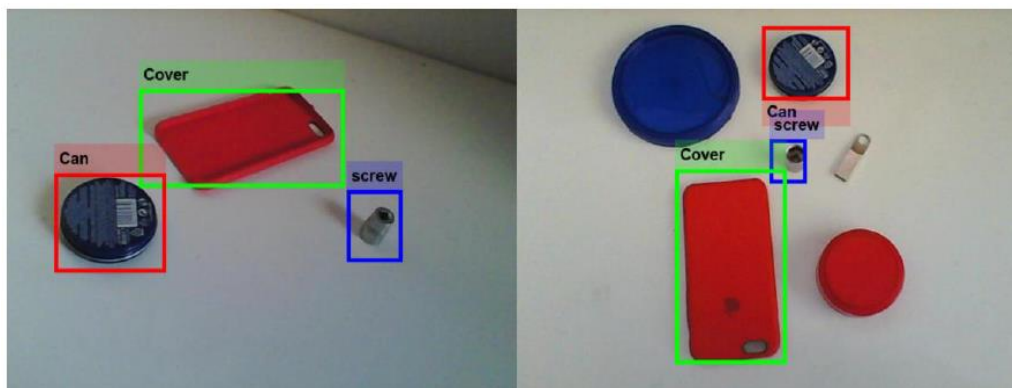
- **Správne pozitívny** – ak bol model identifikovaný a detekcia bola správna
- **Falošne pozitívny** – ak bol model identifikovaný ale detekcia nebola správna

Na určenie, či bol signál správny, alebo falošný bol použitý vzťah pre výpočet IoU (*Intersection over Union*), kde  $P_d$  predstavuje identifikovanú ohraničujúcu plochu a  $P_s$  skutočnú ohraničujúcu plochu. Pri identifikovaní správneho a falošne pozitívneho signálu je požitá prahová hodnota 0,5.

$$IoU = \frac{P_d \cap P_s}{P_d \cup P_s} \quad (2)$$

#### E. Monitorovanie

Posledným krokom učenia siete pre rozpoznávanie objektov bolo jej monitorovanie. V tomto kroku bola naučená sieť podrobená testu, kde sa zisťovala jej schopnosť rozlišovať naučené objekty od podobných objektov. Výsledky tohto testu sú zobrazené na obr.5, kde vľavo sú zobrazené objekty na ktoré bola sieť trénovaná a vpravo rozpoznanie daných objektov medzi ostatnými.



. Obr. 5 Identifikácia naučených objektov



### III. ZÁVER

V tomto článku bol prezentovaný návrh a testovanie algoritmu pre rozpoznávanie objektov s využitím neurónových sietí. Z prezentovaných výsledkov vidno že použitá neurónová sieť dosahuje dostatočnú presnosť pri rozoznávaní objektov na ktoré bola natrénovaná. Obdobný algoritmus je použiteľný aj pri iných aplikáciách kde je potrebné využitie počítačového videnia, napr. v robotických aplikáciách pre polohovanie koncového bodu robotického ramena na základe definovaných značiek prípadne pri klasifikácii, segmentácii alebo detekcii objektov.

### POĎAKOVANIE

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-19-0210

### LITERATÚRA

- [1] E. Šikudová, Z. Černeková, W. Benešová, Z. Haladová, and J. Kučerová, Počítačové videnie Detekcia a rozpoznávanie objektov [online]. Praha: Wikina Praha, 2013. Dostupné na internete: <[http://sccg.sk/~cernekova/Pocitacove\\_videnie.pdf](http://sccg.sk/~cernekova/Pocitacove_videnie.pdf)>
- [2] Ahmed Fawzy Gad, Practical Computer Vision Applications Using Deep Learning with CNNs. New York: Apress media, 2016, ISBN: 978-1-4842-4167-7
- [3] R. Shanmugamani, Deep Learning for Computer Vision. Birmingham: Packt Publishing Ltd, 2018. ISBN: 978-1-78829-562-8
- [4] I. Goodfellow, Y. Bengio, A. Courville, Deep learning [online]. Cambridge: MIT press, Dostupné na internete: <<http://www.deeplearningbook.org/contents/intro.html>>
- [5] P. Sinčák, G. Andrejková, Neurónové siete Inžiniersky prístup(1. diel) [online]. Košice: ELFA, Dostupné na internete: <<http://www.kuzo.szm.com/neuronky1.pdf>>
- [6] Karpathy, A.: Convolutional Neural Networks [Online]. Dostupné na internete: <<https://cs231n.github.io/convolutional-networks/>>
- [7] J. Pavlovičová, J. Polec, M. Patryk, Číslíkové spracovanie obrazu [online]. Bratislava: Vydavateľstvo STU, 2002. Dostupné na internete: <<http://ibooks.sk/publ/12pavlovicova/page1/page1.html>>
- [8] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection [online]. Dostupné na internete: <<https://arxiv.org/pdf/1506.02640.pdf>>
- [9] J. Redmon, A. Farhadi, YOLO9000: Better, Faster, Stronger [online]. Dostupné na internete: <<https://arxiv.org/pdf/1612.08242.pdf>>
- [10] R. Girshick, J. Donahue, T. Darrell, J. Mali, Rich feature hierarchies for accurate object detection and semantic segmentation [online]. Dostupné na internete: <<https://arxiv.org/pdf/1311.2524.pdf>>
- [11] L. Torrey, J. Shavlik, T. Walker, R. Maclin Transfer Learning via Advice Taking [online]. Dostupné na internete: <<https://www.d.umn.edu/~rmaclin/publications/torreyspringer09.pdf>>
- [12] MathWorks, (2020). Deep Learning Toolbox™: User's Guide (R2020a) [Online]. Dostupné na internete: <[https://www.mathworks.com/help/pdf\\_doc/deeplearning/nnet Ug.pdf](https://www.mathworks.com/help/pdf_doc/deeplearning/nnet Ug.pdf)>
- [13] MathWorks, (2020). Computer Vision Toolbox™: User's Guide (R2020a) [Online]. Dostupné na internete: <[https://www.mathworks.com/help/pdf\\_doc/vision/vision Ug.pdf](https://www.mathworks.com/help/pdf_doc/vision/vision Ug.pdf)>
- [14] R. Padilla, Object detection metrics [online]. Dostupné na internete: <<https://github.com/rafaelpadilla/Object-Detection-Metrics>>
- [15] Piotr Skalski, Preventing Deep Neural Network from Overfitting [online]. Dostupné na internete: <<https://towardsdatascience.com/preventing-deep-neural-network-from-overfitting-953458db800a>>
- [16] R. Girshick, Fast R-CNN [online]. Dostupné na internete: <<https://arxiv.org/pdf/1504.08083.pdf>>
- [17] Shaoqing Ren, Kaiming He, R. Girshick, and Jian Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [online]. Dostupné na internete: <<https://arxiv.org/pdf/1506.01497.pdf>>

# Časové charakteristiky prekrývajúcej sa reči v konverzačných interakciách

Stanislav ONDÁŠ, Miriama FEDORIŠINOVÁ

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

<sup>1</sup>stanislav.ondas@tuke.sk, <sup>2</sup>miriama.fedorisinova@student.tuke.sk

**Abstrakt** — Prekrývajúca sa reč je prirodzeným javom, ktorý sa vyskytuje v medziľudských interakciách dialógového typu. Prekrývajúce sa segmenty vznikajú najmä v momentoch striedania sa rečníkov, ako výsledok neúmyselného resp. zámerného prerušenia aktuálneho rečníka. V analyzovanom korpuse nahrávok medziľudských dialógových interakcií sme manuálne označili prekrývajúce sa segmenty za účelom ich ďalšej analýzy. V predloženej práci sme vykonali meranie časových charakteristík týchto segmentov a pokúsili sme sa vyvodit' závery, ktoré by viedli k automatickej klasifikácii komunikačných zámerov jednotlivých prerušení.

**KLúčové slová** — dialóg, prekrývajú sa reč, prerušenia, konverzačná analýza

## I. ÚVOD

Typickým znakom organizácie dialógu je, že účastníci interakcie hovoria jeden po druhom, pričom zmena rečníka ohraničuje jednotlivé prehovory. Tieto zmeny rečníkov, resp. dialógové výmeny sa dejú veľmi plynulo, zvyčajne bez medzery, s veľmi malou medzerou (cca 200ms), resp. s výskytom prekrývajúcej sa reči. Prekrývajúci sa segment reči (overlapping speech) vzniká ako následok situácie, kedy súčasne hovoria viacerí (zvyčajne dvaja) účastníci interakcie. K takejto situácii môže dôjsť neúmyselne, ak dôjde k zlej interpretácii komunikačných signálov aktuálnym poslucháčom, resp. úmyselne, ak sa poslucháč pokúsi prevziať rolu rečníka skôr, ako ten ukončí svoj prehovor. Prekrývajúci sa segment reči vtedy vznikne ako dôsledok úmyselného prerušenia (interruption)[6].

Aj keď v kontexte striedania sa rečníkov považujeme výskyt prekrývajúcej sa reči, resp. skákania do reči ako nežiadúci, zároveň musíme konštatovať, že patrí k základných charakteristikám prirodzenej komunikácie medzi ľuďmi, zvlášť v prípade neformálnych konverzácií, a je potrebné sa ním zaoberať pre komplexné porozumenie interakčných mechanizmov, ktoré sa snažíme preniesť do komunikácie človeka so strojom. Schopnosť byť účastníkom dialógu, resp. konverzácie nie je iba o schopnosti verbalizovať svoje myšlienky a zámery a tiež ich prijímať a spracovávať, ale tiež o schopnosti vnímať, interpretovať a vysielat' komunikačné signály, ktoré umožňujú plynulé a zmysluplné striedanie sa rečníkov, čo v anglickej literatúre označujeme pojmom „turn-taking“ [4], [5].

Interakcie medzi človekom a strojom, kde hlavným komunikačným kanálom je reč a spôsob komunikácie je podobný dialógovej interakcii, sú v prevažnej miere hodnotené ako málo prirodzené. Jedným z dôvodov vnímania neprirodzenosti takýchto interakcií sú aj časové oneskorenia pri striedaní rečníkov. Je možné konštatovať, že zmena rečníka nie je dostatočne plynulá a prirodzená a to najmä z dôvodu, že stroje v úlohe interakčného partnera nedokážu byť dostatočne rýchle, z viacerých príčin. Najpodstatnejšou je fakt, že celý proces spracovania používateľovho vstupu sa začína až v momente, kedy sa tento ukončil, čo je už značne neskoro. Viaceré štúdie (napr. [1], [2], [3]) naznačujú, že ľudský participant pripravuje svoj prehovor už počas prehovoru interakčného partnera, kde v tomto procese využíva aj schopnosť predikovať ešte nevypovedanú výpoveď, čo mu dáva možnosť aj začať svoju repliku skôr, ako svoj prehovor dokončí interakčný partner. Je zrejme, že aj tento fakt, stojí za vznikom prekrývajúcej sa reči. Poznanie mechanizmov a dôvodov vzniku prekrývajúcej sa reči v medziľudských interakciách tak vnímame ako dôležitý poznatok pre vývoj algoritmov, ktoré zrýchlia dialógové výmeny medzi človekom a strojom, prípadne umožnia strojom, v prípade potreby, prevziať rolu rečníka skôr ako dohovori ľudský participant.

Preto sme sa v predloženej práci zamerali na analýzu prekrývajúcich sa segmentov v dyadických dialógových interakciách medzi ľudskými participantmi. Pre tento účel sme pripravili a anotovali korpus dát, na ktorom sme následne označili motivácie jednotlivých prerušení, ktoré viedli k vzniku prekrývajúcich sa segmentov. Následne sme vytvorili automatické nástroje, ktoré nám umožňujú merať časové charakteristiky prekrývajúcich sa častí dialógu.

Kapitola II. opisuje vytvorený korpus dát. V tretej kapitole sú zhrnuté dosiahnuté výsledky spolu s ich analýzou. V závere formulujeme niekoľko hypotéz o časových atribútoch prekrývajúcich sa segmentov v kontexte ich interakčných zámerov, ktoré môžu byť užitočné pre algoritmy striedania sa rečníkov v interakciách medzi človekom a strojom.

## II. KORPUS DÁT

### A. Dáta

Analyzovaný korpus dát pozostával troch rôznych typov nahrávok dialógov – diskusnej relácie *Pod lampou*, dialógov zo seriálu *Oteckovia* a rozhovorov z podcastu *Boris a Brambor*. Dĺžky trvania jednotlivých typov nahrávok sú uvedené v Tab. 1.

Tab. 1  
Korpus dát

Relácia	Počet dielov	Celková dĺžka
Pod lampou	8	07:52:38
<i>Oteckovia</i>	7	00:10:12
<i>Boris a Brambor</i>	2	00:38:00

Analyzovaný korpus má celkové trvanie približne 8 hodín a 40 minút a vznikol rozšírením korpusu použitého v [7]. Počas tohto času došlo k 1693 dialógovým výmenám (zmene rečníka), pričom celkový počet segmentov prekrývajúcej sa reči bol 884.

Nový korpus dát vznikol rozšírením pôvodného korpusu, ktorý obsahoval iba epizódy relácie *Pod lampou*. V prípade tejto relácie sa jedná o moderovanú diskusiu, resp. interview, čo predstavuje iba jednu skupinu dialógových scenárov, navyše s fixnými rolami jednotlivých participantov. Pre interview, resp. televízne diskusie je typické, že jeden účastník má rolu moderátora a druhý účastník je v postavení respondenta. Je zrejmé, že tieto role nie sú rovnocenné, čo sa zvlášť prejavuje v miere, akou sa jednotliví účastníci podieľajú na manažovaní interakcie. Oveľa zaujímavejšou skupinou sú rozhovory dvoch rovnocenných účastníkov, pričom ich role sú viac vyvážené. Preto sme k relácii *Pod lampou* hľadali dialógy, ktoré by sa viac podobali prirodzeným konverzáciám. Za veľmi prirodzenú interakciu považujeme dialógy v denných seriáloch, ktoré aj napriek tomu, že sú hrané, dosahujú vysokú mieru autentickosti. Druhým typom relácie je podcast *Boris a Brambor*, v ktorom sme identifikovali úseky dialógov, kde prebieha dialóg iba medzi oboma moderátormi, kde obaja hrajú rolu diskutujúcich.

Rôznorodosť analyzovaného korpusu nám umožnila pozorovať a popísať rozdiely v striedaní rečníkov v rôznych typoch dyadických interakcií.

### B. Anotačná schéma

Na označenie prehovorov rečníkov a prekrývajúcich sa segmentov bol použitý anotačný nástroj Anvil, pre ktorý sme vytvorili anotačnú schému pozostávajúcu z dvoch častí. Prvá časť (Obr. 1) umožňuje anotovať jednotlivé prehovory rečníkov.

```
<track-spec name="Event_s2" type="primary">  
  <attribute name="type" valueType="overlapType" />  
</track-spec>  
  
<track-spec name="Turn_1" type="primary">  
  <attribute name="type" valueType="String" />  
</track-spec>
```

Obr. 1 Časť anotačnej schémy pre označovanie prehovorov rečníkov

Druhá časť anotačnej schémy (Obr.2) je určená pre označovanie prekrývajúcich sa

segmentov a ich prislúchajúcich interakčných zámerov, resp. motivácií.

```
<valueset name="overlapType">
    <value-el>obtain_confirmation</value-el>
    <value-el>obtain_clarification</value-el>
    <value-el>obtain_agreement</value-el>
    <value-el>provide_additional_info</value-el>
    <value-el>complete_information</value-el>
    <value-el>anticipation</value-el>
    <value-el>support_speaker</value-el>
    <value-el>express_disagreement</value-el>
    <value-el>express_aggrement</value-el>
    <value-el>express_involvement</value-el>
    <value-el>express_noninterest__topic_shift</value-el>
    <value-el>try_to_take_floor</value-el>
    <value-el>turn_leaving_misunderstanding</value-el>
    <value-el>stop_speaker</value-el>
    <value-el>additional_question</value-el>
    <value-el>unclear</value-el>
</valueset>
```

Obr. 2. Časť anotačnej schémy pre označenie prekrývajúcich sa segmentov a ich označení

Zoznam interakčných zámerov prerušujúcich prehovorov, ktoré vytvorili prekrývajúce sa časti v dialógu sú zhrnuté v Tab. 2

Tab. 2  
 Interakčné zámery prerušení

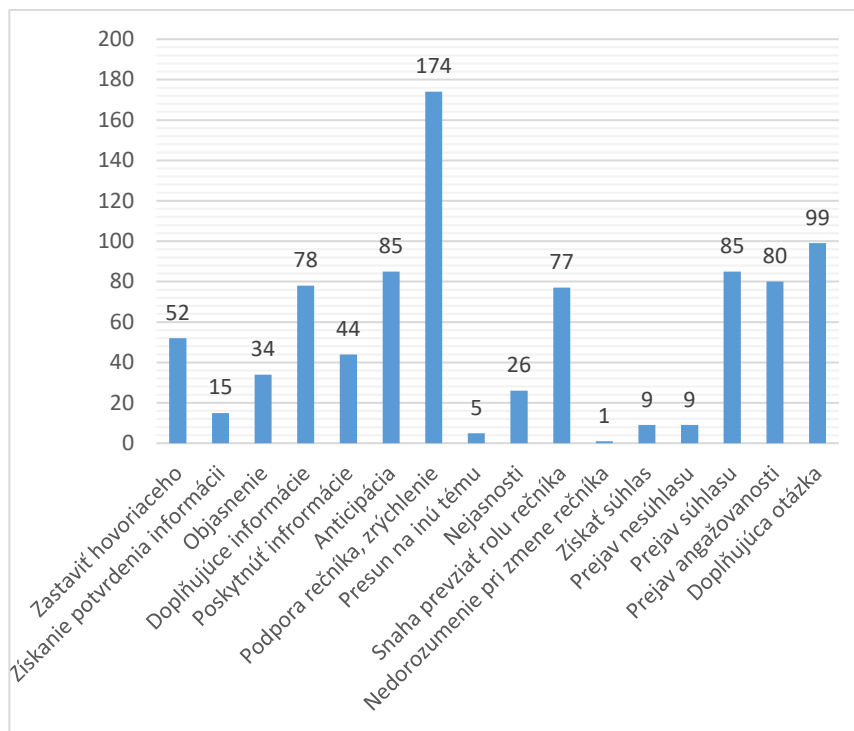
Prejav súhlasu ( <i>Express agreement</i> )	Doplnenie informácie ( <i>Complete information</i> )
Prejav nesúhlasu ( <i>Express disagreement</i> )	Poskytnutie dodatočnej informácie ( <i>Provide additional info</i> )
Prejav angažovanosti ( <i>Express involvement</i> )	Doplňujúca otázka ( <i>Additional question</i> )
Vyjadrenie nezájmu, presun na inú tému ( <i>Express noninterest topic shift</i> )	Podpora rečníka, zrýchlenie interkcie ( <i>Support speaker</i> )
Získanie/vyžiadanie súhlasu ( <i>Obtain agreement</i> )	Zastaviť hovoriaceho ( <i>Stop speaker</i> )
Získanie potvrdenia informácií ( <i>Obtain confirmation</i> )	Snaha prevziať rolu rečníka ( <i>Try to take floor</i> )
Získať vysvetlenie ( <i>Obtain clarification</i> )	Nedorozumenie pri zmene rečníka ( <i>Turn leaving misunderstanding</i> )
Nejasný dôvod prerušenia ( <i>Unclear interruption</i> )	Predikcia ( <i>Anticipation</i> )

### III. VÝSLEDKY

Na analyzovanom korpuse nahrávok sme vyhodnotili početnosti interakčných zámerov jednotlivých prerušení, ktoré spôsobili vznik prekrývania sa reči. Najpočetnejšou kategóriou interakčných zámerov sa stal zámer „Support speaker” (podpora rečníka, zrýchlenie interakcie). Tento komunikačný zámer veľmi často využívajú moderátori diskusií resp. interview, pri manažovaní interakcie, tak, aby jednak pomohli respondentovi vyjadriť jeho myšlienky a postoje a tiež aby zabezpečili napredovanie interakcie, resp. manažovanie času určeného pre danú reláciu/interview. Za zvýšenú početnosť tohto zámeru môže najmä fakt, že prevažná časť korpusu pozostáva práve z relácií tohto typu (Pod lampou).

Druhou najpočetnejšou a zároveň veľmi vyváženou skupinou boli interakčné zámery, ktoré súvisia s vyjadrením pozornosti, angažovanosti a záujmu o interakciu – prejav súhlasu, prejav angažovanosti, poskytnutie doplňujúcej informácie, polozenie doplňujúcej otázky. V prípade týchto zámerov môžeme konštatovať, že reprezentujú kooperačné činnosti vykonávané v rámci interakcie, ktorých cieľom je byť aktívnym participantom interakcie so snahou, aby táto bola úspešná.

Tretia skupina interakčných zámerov patrí do skupiny tzv. konkurenčných prerušení, kde je zámerom poslucháča prerušiť aktuálneho rečníka, resp. prevziať rolu rečníka na seba.



Obr. 1. Vyskyt interakčných zámerov prerušení v analyzovanom korpuse

Pre pozorovanie časových charakteristík prekrývání sme vytvorili jednoduché automatické nástroje v jazyku Python, schopné spracovať výstupné súbory anotačného nástroja Anvil a z nich extrahovať údaje o dĺžkach trvania jednotlivých prehovorov a prekrývajúcich sa častí.

V nasledujúcej tabuľke sme vyhodnotili priemerné dĺžky jednotlivých typov prekrývajúcich sa segmentov.

Tab. 3  
Priemerné dĺžky prekrývajúcich sa segmentov reči podľa interakčného zámeru [s]

Motivácia k prerušeniu	Pod lampou	Podcasty	Oteckovia
Zastaviť hovoriaceho	1,23	0,79	0,66
Získať vysvetlenie	1,79	2,56	0,48
Doplnenie informácie	1,86	1,15	0,48
Snaha prevziať rolu rečníka	1,48	2,1	0,87
Prejav súhlasu	0,89	0,89	0,92
Prejav nesúhlasu	1,91	1,16	2,6
Prejav angažovanosti	1,52	1,5	0,56
Doplňujúca otázka	1,69	1,00	1,59

Namerané priemerné dĺžky prekrývajúcich sa segmentov ukázali rozdiely najmä medzi konverzáciami v relácii Pod lampou a v seriáli Oteckovia, kde môžeme pozorovať výrazne nižšie dĺžky prekrývajúcich sa segmentov, čo môže byť spôsobené celkovo vyššou dynamikou dialógových výmen v dennom seriáli. Naopak, v prípade prejavu súhlasu je možné pozorovať približne zhodné dĺžky trvania (0,89s – 0,92s) prekrývajúcich sa segmentov, čo nasvedčuje využitiu rovnakých interakčných a verbálnych prostriedkov pri tomto zámere. Zároveň je možné pozorovať výrazný rozdiel medzi prejavom súhlasu a prejavom nesúhlasu, kde prejavy nesúhlasu trvajú signifikantne dlhšie, čo naznačuje väčší dôraz poslucháča na to, aby tento prejav bol vnímaný a akceptovaný. V ostatných kategóriách je zložité vyvodit' konkrétne závery vzhľadom na malý objem dát.

#### IV. ZÁVER

Predložená práca sa venuje analýze prekrývajúcich sa segmentov reči v konverzačných

medziľudských interakciách, ako dôležitému fenoménu pri striedaní sa rečníkov. Na takmer deväťhodinovom korpuse nahrávok dialógov sme identifikovali a označili prekrývajúce sa segmenty a priradili sme im najpravdepodobnejší interakčný zámer osoby, ktorá prerušila aktuálneho rečníka. Následne sme vyhodnotili dĺžky trvania jednotlivých prekrývajúcich sa segmentov v kontexte ich interakčných zámerov. Dosiahnuté výsledky naznačujú, že väčšina prekrývajúcich sa segmentov netrvá viac ako dve sekundy. Ďalším zistením je, že dĺžka prekrývajúcich sa segmentov pravdepodobne závisí aj od typu a dynamiky konverzácie. Je zrejmé, že táto oblasť si vyžaduje ďalšie skúmanie.

Výsledky pozorovania dĺžok prekrývajúcich sa segmentov by mali byť nápomocné pri návrhu systémov s rečovou interakciou medzi človekom a strojom, kde by mohli byť aplikované na estimáciu vhodného času pre načasovanie reakcie stroja na repliku ľudského účastníka dialógu. Zároveň, táto práca predstavuje materiál na ďalšie skúmanie časových charakteristík striedania sa rečníkov, najmä s ohľadom na skúmanie času, kedy aktuálny poslucháč začína plánovať svoj vlastný prehovor.

#### POĎAKOVANIE

Táto práca vznikla s podporou projektov APVV SK-TW-2017-0005, KEGA 009TUKE-4/2019, VEGA 1/0753/20 a VEGA 2/0165/21 financovaných z prostriedkov Ministerstva školstva, vedy, výskumu a športu SR.

#### LITERATÚRA

- [1] L. Magyari, M.C. Bastiaansen, J.P. de Ruiter, and S.C. Levinson, "Early Anticipation Lies behind the Speed of Response in Conversation", *Journal of Cognitive Neuroscience*, pp. 2530-2539, 2014.
- [2] R.S. Gisladottir, S. Bögels, and S.C. Levinson, "Oscillatory Brain Responses Reflect Anticipation during Comprehension of Speech Acts in Spoken Dialog", *Front. Hum. Neurosci.*, 2018.
- [3] A.J. Liddicoat, "The projectability of turn constructional units and the role of prediction in listening", *Discourse Studies*, Vol.6, No.4, pp. 449-469, 2004.
- [4] J. Holler, K. H. Kendrick, M. Casillas, S. C. Levinson, eds. (2016). *Turn-Taking in Human Communicative Interaction*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-825-2, 2016
- [5] K. R. Thórisson, "Natural turn-taking needs no manual: computational theory and model, from perception to action," in *Multimodality in Language and Speech Systems*, eds B. Granström, D. House, and I. Karlsson (Netherlands: Springer), 173–207, 2002
- [6] M. Heldner, and J. Edlund, "Pauses, gaps and overlaps in conversations." *J.Phon.* 38, 555–568, 2010
- [7] S. Ondáš, M. Pleva, „Anticipation and its applications in human-machine interaction“, in: *ITAT 2019*, vol. 2473: Proceedings of the 19th conference Information Technologies - Applications and Theory, CEUR-WS, Nemecko, s. 152-156, - ISSN 1613-0073

# Riadenie horizontálneho zásobníka kontinuálnej linky

<sup>1</sup>Lukáš KONIAR, <sup>2</sup>František ĎUROVSKÝ

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

<sup>1</sup>lukas.koniar@student.tuke.sk, <sup>2</sup>frantisek.durovsky@tuke.sk

**Abstrakt** — Článok sa zaoberá riadením výstupného horizontálneho zásobníka materiálu na kontinuálnej linky. Obsahuje matematický popis zásobníka, návrh jeho riadenia a analýzu vplyvu jednotlivých veličín na ťah v materiáli. Analýza je robená simulačne v programovom prostredí Matlab/Simulink. Parametre zásobníka použité v analýze zodpovedajú reálnej linky na spracovanie oceľových plechov.

**KLúčové slová** — horizontálny zásobník materiálu, kontinuálna linka, matematický model, MATLAB, Simulink

## I. ÚVOD

V kontinuálnych linkách na spracovanie materiálu je nutné zabezpečiť, aby cez technologickú časť prechádzal materiál bez prerušenia. Jedná sa o linky, kde materiál prechádza cez pece (žíhacie linky), kyselinový roztok (moriace linky), elektrolytické vane (pocínovanie) alebo roztavený kov (pozinkovacie linky). Materiál sa do linky zvyčajne odoberá zo zvitku a na konci linky sa navíja na iný zvitok. Zavádzanie materiálu do takýchto liniek je komplikované. Pri zavádzaní materiál na začiatku a konci pásu nie je spracovávaný za štandardných podmienok (nastavené ťahy, teploty a rýchlosť pásu), takže nemá požadovanú kvalitu. Aby sa kvôli zavádzaniu pásu technologický proces nemusel prerušovať, pás z nového zvitku sa na vstupe do linky privarí na koniec predchádzajúceho pásu. Počas manipulácie s pásom na vstupe (výmena zvitku, zarovnanie začiatku pásu, zvarenie pásov, obrúsenie zvaru), pás na vstupe linky stojí. Technologická časť zatiaľ spracúva pás, ktorý odoberá z tzv. vstupného zásobníka. Keď sa pásy spoja a upraví, vstupná časť linky sa rozbehne na rýchlosť vyššiu ako má technologická časť, a materiál sa vo vstupnom zásobníku doplní na maximum. Vstup linky je tak pripravený na nasledujúcu výmenu zvitku. Podobná situácia je aj na výstupe linky. Tu sa počas výmeny zvitku na navíjačke materiál z technologickú časť ukladá vo výstupnom zásobníku.

Zásobníky materiálu môžu byť horizontálne alebo vertikálne. Keďže vzdialenosti medzi valcami v naplnenom zásobníku sú veľké, je nutné pri opise zásobníka uvažovať aj s dĺžkou, pružnosťou príp. hmotnosťou pásu. Situáciu komplikuje aj fakt, že pásy materiálu predstavujú pružné členy s premenlivými parametrami [1].

Článok rieši návrh riadenia pre horizontálny zásobník oceľového pásu a analýzu vplyvu jednotlivých veličín, ako sú trenie a dynamická kompenzácia, na presnosť ťahu v oceľovom páse. Analýza bola robená simulačne s využitím parametrov reálnej linky.

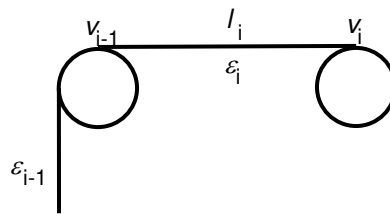
## II. MATEMATICKÝ MODEL HORIZONTÁLNEHO ZÁSOBNÍKA MATERIÁLU

Matematický model horizontálneho zásobníka možno rozdeliť na dve časti: model pásu v zásobníku a model slučkového vozíka, ktorý pás napína. Model pružného pásu medzi dvomi valcami vychádza z Obr. 1. Ak zanedbáme šmykanie pásu po valcoch, rovnica kontinuity aplikovaná na transport pásu bude [2], [3]

$$\frac{d}{dt} \left( \frac{l_i}{1 + \varepsilon_i} \right) = -\frac{v_i}{1 + \varepsilon_i} + \frac{v_{i-1}}{1 + \varepsilon_{i-1}}, \quad (1)$$

kde  $l_i$  je vzdialenosť medzi valcami,  $v_i$  je obvodová rýchlosť na valci a  $\varepsilon_i$  je predĺženie pásu na valci. Pre malé hodnoty predĺženia platí, že

$$\frac{1}{1+\varepsilon} \approx 1-\varepsilon \quad (2)$$



Obr. 1 Dva valce spojené pružným pásom

Rovnicu (1) potom môžeme prepísať na tvar:

$$\frac{d}{dt} [l_i(t)(1-\varepsilon_i(t))] = v_{i-1}(1-\varepsilon_{i-1}) - v_i(1-\varepsilon_i) \quad (3)$$

a po úprave (3) dostaneme

$$l_i(t) \frac{d\varepsilon_i(t)}{dt} = v_i(1-\varepsilon_i) - v_{i-1}(1-\varepsilon_{i-1}) + [1-\varepsilon_i(t)] \frac{dl_i(t)}{dt} \quad (4)$$

Podľa Hookovho zákona platí

$$\varepsilon = \frac{\sigma}{E} \quad (5)$$

kde  $\sigma$  je napätie v páse ( $\text{Pa/m}^2$ ) a  $E$  je Youngov modul pružnosti. Predĺženie  $\varepsilon$  je dané ako

$$\varepsilon = \frac{l_1 - l_0}{l_1} \quad [-] \quad (6)$$

Sila v páse je definovaná ako

$$F = S\sigma = SE\varepsilon \quad \text{a z toho} \quad \varepsilon = \frac{F}{SE} \quad (7)$$

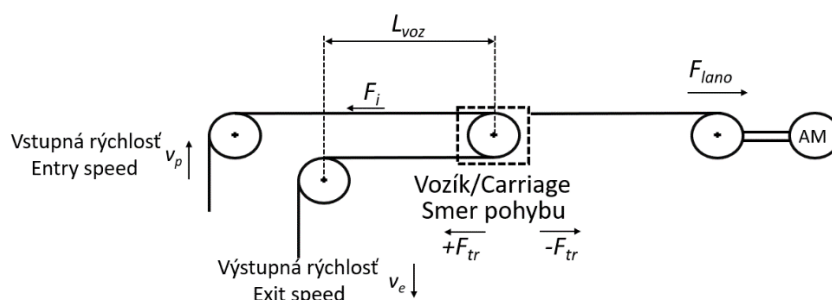
kde  $S$  je prierez materiálu a  $F$  je sila pôsobiaca na pás. Dosadíme (7) do (4) a upravíme:

$$\frac{d}{dt} (F_i(t)) = \frac{SE}{l_i(t)} (v_i - v_{i-1}) + \frac{1}{l_i(t)} (v_{i-1}F_{i-1}(t) - v_iF_i(t)) + \frac{SE}{l_i(t)} \frac{dl_i(t)}{dt} - \frac{F_i(t)}{l_i(t)} \frac{dl_i(t)}{dt} \quad (8)$$

Pri oceľových pásoch je Youngov modul pružnosti  $210 \cdot 10^9$  Pa, takže druhý člen na pravej strane rovnice (8) bude rádovo menší, preto ho v ďalších výpočtoch zanedbáme. Výsledná rovnica pre pás s premenlivou dĺžkou potom bude

$$\frac{d}{dt} (F_i(t)) = \frac{SE}{l_i(t)} (v_i - v_{i-1}) + \frac{SE}{l_i(t)} \frac{dl_i(t)}{dt} - \frac{F_i(t)}{l_i(t)} \frac{dl_i(t)}{dt} \quad (8)$$

Pri horizontálnom zásobníku sa materiál ukladá vo vrstvách nad sebou. Konštrukcia zásobníka nevyžaduje vysokú halu, na druhej strane vrstvy materiálu vytvárajú kvôli vlastnej hmotnosti previsy, ktoré je nutné podoprieť, napr. separátormi, aby nedošlo k vzájomnému poškrabaniu materiálu. Vozík s valcami sa pohybuje po vodorovnej dráhe na kofajniciach a v závislosti od požadovaného množstva akumulovaného materiálu a dĺžky dráhy, ktorá je k dispozícii, môže byť na vozíku jeden alebo viac vratných valcov. V našom prípade sa jednalo o vozík s jedným vratným valcom. Vozík bol do jednej strany ťahaný lanovým navijakom a do druhej strany samotným pásom. Usporiadanie zásobníka je na Obr. 2.



Obr. 2 Usporiadanie horizontálneho zásobníka materiálu



Rovnica (8) opisuje dynamiku sily v jednom páse. Pri podrobnej analýze tak umožňuje zahrnúť do modelu aj moment zotrvačnosti vratného valca na vozíku. Pri zjednodušenej analýze budeme predpokladať, že vratný valec neovplyvňuje výrazne dynamiku pásu, takže ho do analýzy nezahrnieme. V tom prípade bude sila v hornom aj dolnom páse rovnaká. Rovnica pre dynamiku celkovej sily, ktorou bude pás pôsobiť proti navijaku, bude podľa Obr. 2:

$$\frac{d}{dt}(F_i(t)) = \frac{SE}{L_{voz}(t)N}(v_e - v_p) + \frac{SE}{L_{voz}(t)} \frac{dL_{voz}(t)}{dt} - \frac{F_i(t)}{L_{voz}(t)} \frac{dL_{voz}(t)}{dt}, \quad (9)$$

kde  $L_{voz}$  je okamžitá vzdialenosť vozíka od parkovacej polohy,  $N$  je počet pásov v zásobníku (v našom prípade  $N=2$ ),  $v_p$  je vstupná a  $v_e$  výstupná rýchlosť pásu do/zo zásobníka.

Vozík s vratnými valcami s celkovou hmotnosťou  $m_{voz}$  je zvyčajne ťahaný lanom. Pri naplňaní zásobníka je vozík ťahaný lanom silou  $F_{lano}$ . Proti tejto sile pôsobí ťahová sila v materiáli  $F_{tah}$ , daná súčtom ťahových síl v jednotlivých pásoch, a straty trením pohybujúceho sa vozíka po koľajniciach  $F_{tr}$ . Rozdiel medzi týmito silami je sila  $F_{dyn}$ , ktorá spôsobuje zrýchlenie resp. spomalenie vozíka  $a_{voz}$ . Pre silu v lane podľa [4] platí:

$$F_{lano}(t) = F_{tah}(t) + F_{tr}(t) + F_{dyn} = \sum_{i=1}^N F_i(t) + F_{tr}(t) + m_{voz} a_{voz}(t). \quad (10)$$

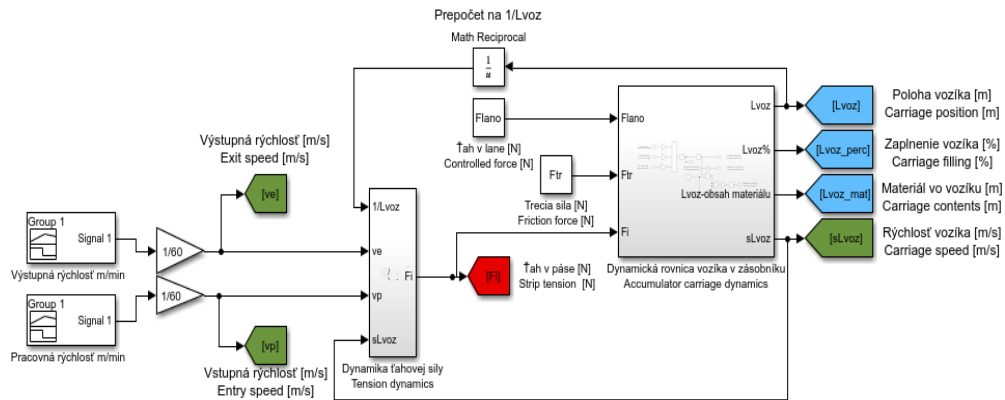
Zrýchlenie vozíka vyjadríme pomocou vzdialenosti  $L_{voz}$

$$a_{voz}(t) = \frac{d^2 L_{voz}(t)}{dt^2}. \quad (11)$$

Rovnicu (9) s uvážením (11) upravíme a dostaneme dynamickú rovnicu pre pohyb vozíka

$$\frac{d^2 L_{voz}(t)}{dt^2} = \frac{1}{m_{voz}} F_{lano}(t) + \frac{1}{m_{voz}} F_{tr}(t) + \frac{1}{m_{voz}} \sum_{i=1}^N F_i(t). \quad (12)$$

Simulačné schéma celého horizontálneho zásobníka materiálu je na Obr. 3.



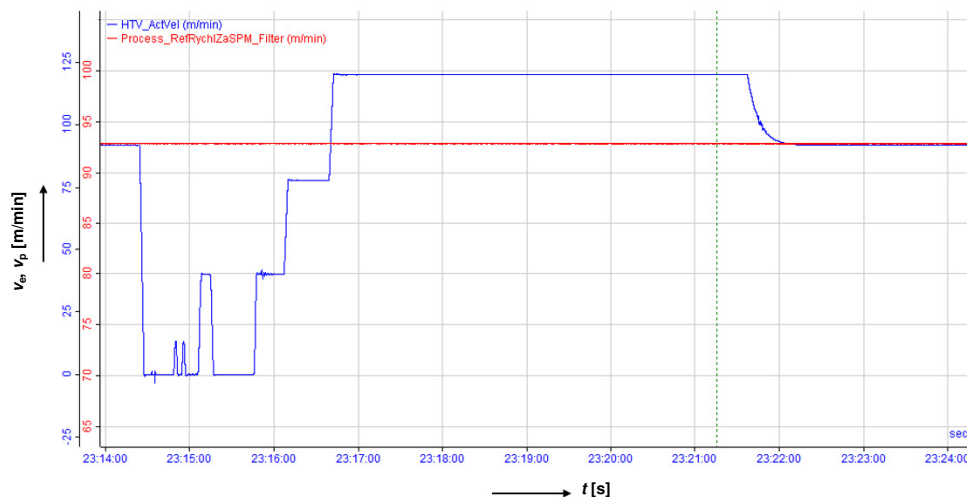
Obr. 3 Simulačné schéma výstupného horizontálneho zásobníka

### III. VPLYV PARAMETROV NA ŤAH V PÁSE V ZÁSObNÍKU

Horizontálny zásobník materiálu je zariadenie, ktoré obsahuje viacero pružných členov: dve vrstvy ukladaného pásu a napínacie lano z navijaku. Ťahy v páse ani v napínacom lane nie sú merané, preto je nutné ich riadiť nepriamo. Jedinou informáciou o ťahových pomeroch v zásobníku je moment motora navijaku. Ten obsahuje zložku na vytvorenie ťahu v lane, kompenzáciu trenia vozíka na koľajniciach a dynamický moment pri zmenách rýchlosti vozíka. Nepresné nastavenie uvedených zložiek vedie ku zmenám ťahu v páse, čo sa nepriaznivo prejavuje na ťahoch a rýchlostiach celého výstupného úseku. Kritická situácia je počas prechodných dejov, kedy sa na výstupnom úseku odoberá navinutý zvitok a pás sa zavádza do nového zvitku. Reálne rýchlostné pomery výstupného úseku pri výmene zvitku sú na Obr. 4.

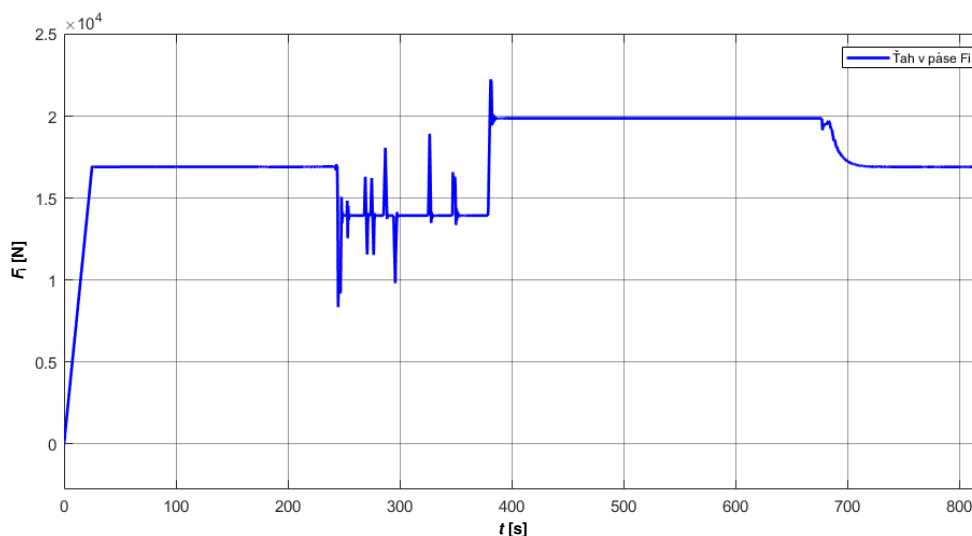
Na Obr. 4 je znázornená rýchlosť výstupného úseku (v obrázku je označená ako HTV\_Actvel) a predstavuje výstupnú rýchlosť zásobníka materiálu  $v_e$ . Výstupný úsek sa po navinutí potrebného množstva materiálu zastaví a pás sa v nožniciach pred navijackou prestrihne. Navijacka dovinie pás a zvitok je z nej odobraný. Otáčky navijacky sú po odstrihnutí pásu iné ako rýchlosť výstupného úseku, takže sa na Obr. 4 neprejaví. Následne sa pás dvakrát posunie pod nožnice na zarovnanie začiatku, odstrihnutie odpadu príp. odobranie vzoriek pre laboratórium. Potom je zavedený do navijacky a pripravený na začiatok navijania. Následne sa rozbieha spolu s navijackou, je navijackou letmo uchytený, a ďalej sa rýchlosť výstupného úseku po krokoch zvyšuje na 120 m/min. Pri tejto rýchlosti sa zásobník vyprázdni a po

vyprázdnení rýchlosť výstupného úseku klesne na rýchlosť procesného úseku. Kým je rýchlosť výstupného úseku nižšia ako procesná, zásobník sa naplňuje, keď je vyššia, tak sa vyprázdňuje.



Obr. 4 Rýchlosť výstupného úseku (modrá) a procesného úseku (červená).

Simulovaný priebeh ťahu v páse v zásobníku počas výmeny zvitku je na Obr. 5. Voči reálnemu stavu boli v simulácii urobené určité úpravy. Aby sme na začiatku simulácie dosiahli ustálené pomery zodpovedajúce ustálenému pracovnému stavu na linke, je nutné pás najprv napnúť a linku rozbehnúť na pracovné hodnoty. Preto ťah a rýchlosť pásu i linky na začiatku simulácie rástli po rampe na pracovné hodnoty. Skoková zmena by spôsobila rozkmitanie pásu. Analýza priebehov zo simulácie sa preto bude robiť po ustálení ťahov a rýchlostí, čo predstavuje čas od cca 100 s ďalej. Ťah v páse v zásobníku na Obr. 5 je zobrazený pre prípad riadenia bez kompenzácií, len s nastaveným konštantným ťahovým momentom navijaku. Hodnota nastaveného ťahu je cca 17 kN, čo je dané technologickým predpisom pre daný spracovávaný sortiment. Túto hodnotu systém udržiava v ustálenom stave. Pri naplňaní zásobníka sa časť momentu motora navijaku spotrebuje na prekonanie trenia vozíka na koľajniciach, čo spôsobí pokles ťahu v páse. Pri vyprázdňovaní zásobníka je naopak vozík ťahaný pásom, navijak ho len brzdí. Tretí moment je v tomto prípade krytý pohonmi na vstupe a výstupe zásobníka, a cez pás sa prenáša na vozík. To spôsobí nárast ťahu v páse. Pri zmenách rýchlosti výstupného úseku sa na ťahu prejaví nevykompenzovaný dynamický moment potrebný na urýchlenie resp. spomalenie vozíka zásobníka, ktorý spôsobí krátkodobé zmeny ťahu – špičky na Obr. 5. Posunutie priebehov v čase voči Obr. 4 je spôsobené rôznymi mierkami meraných a simulovaných priebehov.

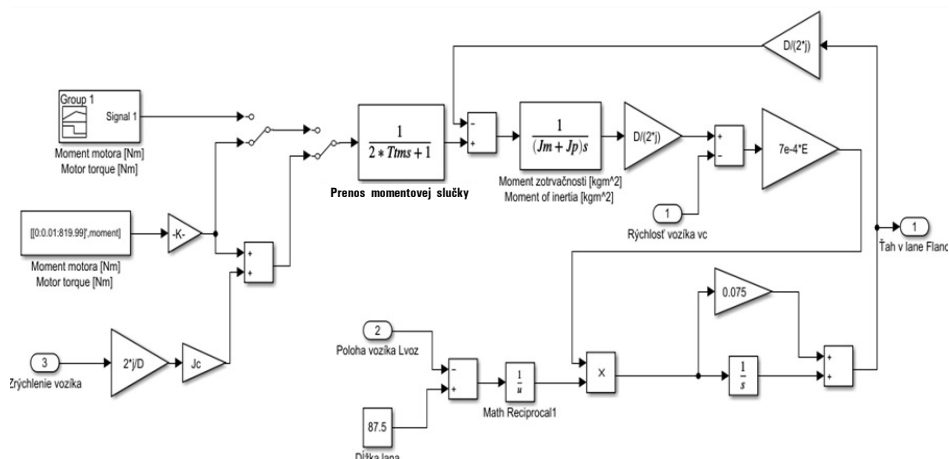


Obr. 5 Priebeh ťahovej sily v páse zásobníka pri konštantnej ťahovej sile navijaku bez kompenzácií

#### IV. RIADENIE VOZÍKA HORIZONTÁLNEHO ZÁSOBNÍKA

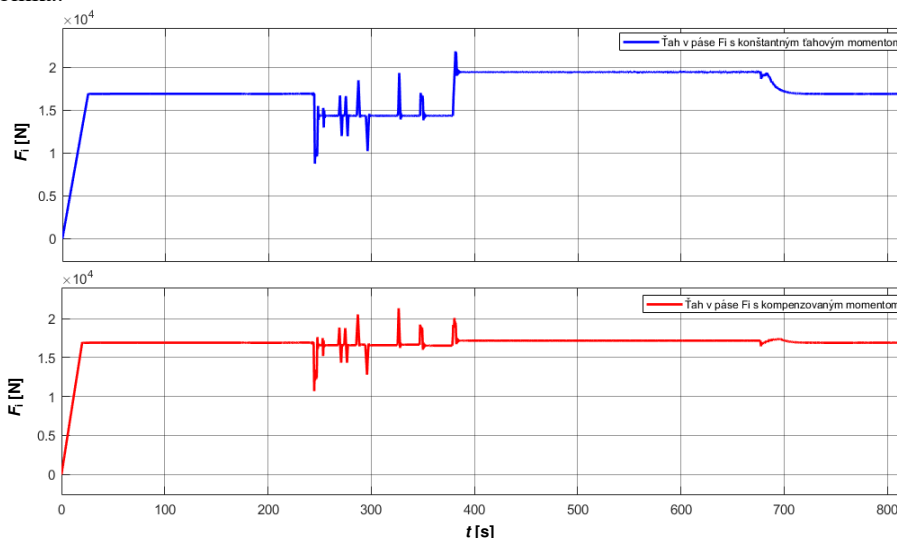
Navijak výstupného zásobníka je poháňaný cez prevodovku asynchrónnym motorom napájaným z frekvenčného meniča [5]. Motor pracuje v pásme do menovitých otáčok, takže nie je nutné jeho odbudzovanie. Riadenie je nastavené ako momentové [6]. Simulačná schéma

pohonu je na Obr. 6. Riadenie meniča generuje želaný moment, ktorý podradená regulácia momentotvornej zložky prúdu zabezpečí bez ohľadu na to, aké je budenie stroja. Preto pri simulácii stačí momentovú slučku nahradiť zotrvačnosťou 1.rádu [7]. Na Obr. 6 je v pravej dolnej časti znázornený aj model pružného lana.



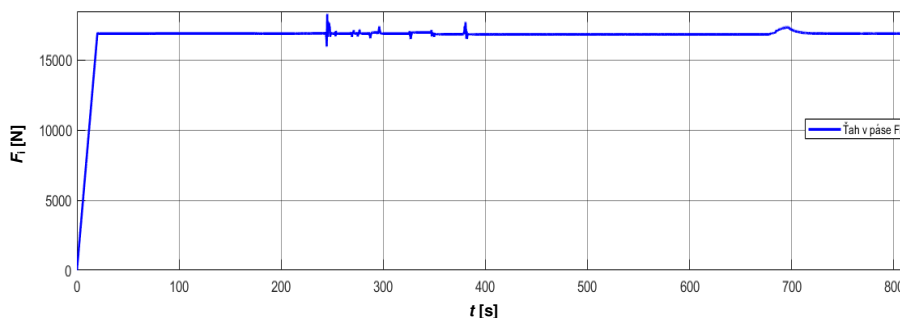
Obr. 6 Simulačná schéma riadenia vozíka výstupného horizontálneho zásobníka

Keďže nie je k dispozícii informácia o ťahu v páse v zásobníku, riadenie ťahu môže byť len nepriame. Ťah je riadený prostredníctvom momentu motora. Nepriame riadenie ťahu pracuje v otvorenej regulačnej slučke, preto vyžaduje čo najpresnejšiu kompenzáciu trecieho a dynamického momentu. Ich nepresná kompenzácia sa prejaví na ťahu tak, ako to vidno na Obr. 5 a v hornej časti Obr. 7. V dolnej časti Obr. 7 je priebeh ťahu s vykompenzovaným trecím momentom. Špičky v ťahu zostali kvôli nevykompenzovanému dynamickému momentu vozíka zásobníka.



Obr. 7 Porovnanie ťahov v páse pre konštantný ťahový (hore) a kompenzovaný trecí moment (dole)

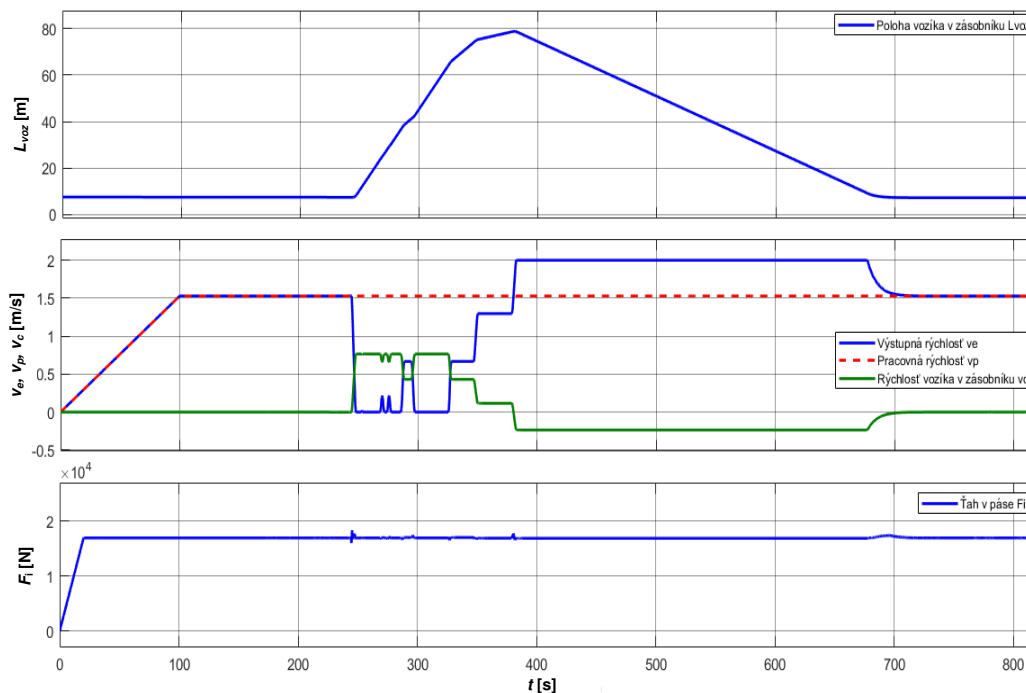
Ak ku želanému momentu motora pridáme čiastočnú kompenzáciu dynamického momentu vozíka zásobníka, dostaneme priebeh ťahu podľa Obr. 8.



Obr. 8 Ťah v páse s vykompenzovaným trecím a čiastočne aj dynamickým momentom

Na Obr. 8 vidno zlepšenie priebehu ťahovej sily, ale priebeh stále vykazuje zakmitanie sily. Pri presnej kompenzácii dynamického momentu dostaneme priebeh podľa Obr. 9 (dole), kde je zakmitanie ťahu pri zmenách rýchlosti vozíka minimálne. Veľkosť trecej sily ako aj moment zotrvačnosti vozíka boli identifikované na reálnej linke a následne využité v simulácii.

Obr. 9 ukazuje aj priebehy ostatných dôležitých veličín počas výmeny zvitku na výstupe linky: polohu slučkového vozíka (hore), rýchlosti výstupného úseku (S-valce, navíjačka) – modrá farba, rýchlosti slučkového vozíka – zelená farba a rýchlosti procesného úseku – červená farba (v strede), a ťahu v páse (dole). Z priebehu ťahu v páse je evidentné, že presná kompenzácia dynamického a trecieho momentu výrazne zlepšuje presnosť riadenia ťahovej sily výstupného zásobníka. Viac podrobností a výsledkov je uvedených v [8].



Obr. 9 Priebehy hlavných veličín výstupného úseku vrátane horizontálneho zásobníka.

## V. ZÁVER

V článku je prezentované riadenie horizontálneho zásobníka materiálu na kontinuálnej linke. Bol zostavený matematický model zásobníka, ktorý obsahuje dynamické modely pásu v zásobníku, napínacieho lana navíjaku a pohonu navíjaku. Analýza riadenia bola vykonaná simulačne a ukázala vplyv trenia medzi vozíkom a pojazďovými koľajnicami a momentu zotrvačnosti slučkového vozíka na presnosť ťahu v páse. Analýza bola robená simulačne s využitím parametrov reálnej výrobnéj linky.

## POĎAKOVANIE

Táto práca bola podporovaná Vedeckou grantovou agentúrou Ministerstva školstva Slovenskej republiky v rámci projektu VEGA 1/0493/19.

## LITERATÚRA

- [1] G. Brandenburg, „Ein mathematisches Modell für eine durchlaufende elastische Stoffbahn in einem System angetriebener, umschlungener Walzen“, *Regelungstechnik und Prozess-Datenverarbeitung*, vol. 3, pp. 69–162, 1973.
- [2] P. R. Pagilla, I. Singh, R. V. Dwivedula, „A Study on Control of Accumulators in Continuous Web Processing Lines“, in *Proc. of American Control Conference*, Denver, Colorado, p.3687 – 3689, 2003.
- [3] H. Koç, D. Knittel, G. Abba, M. de Mathelin, C. Gauthier, and E. Ostertag, “Modeling and control of an industrial accumulator in a web transport system,” in *Proc. Europ. Contr. Conf. ECC’99*, 1999.
- [4] V. Fedák, *Modelovanie elektromechanických sústav*. Technická univerzita v Košiciach, 2015, 243 s. ISBN 978-80-553-2111-0
- [5] L. Zboray, F. Ďurovský, J. Tomko, *Regulované pohony*, Viena Košice, 2000. ISBN 80-88922-13-5.
- [6] F. Ďurovský, K. Kyslan, *Prednášky z predmetu Servopohony – Regulačné štruktúry pohonov*. Technická univerzita v Košiciach, 2019.
- [7] K. Kyslan, F. Ďurovský, *Regulované pohony – Návod na cvičenia*. Košice. Technická univerzita v Košiciach, 2016. 68s. ISBN 978-80-553-2627-6
- [8] L. Koniar, „Riadenie horizontálneho zásobníka kontinuálnej linky“. Diplomová práca FEI TU Košice, 2021.

# Overovanie hodnovernosti biometrického znaku

<sup>1</sup>Eduard PYNZAR, <sup>2</sup>Matúš PLEVA

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

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

<sup>1</sup>eduard.pynzar@student.tuke.sk, <sup>2</sup>matus.pleva@tuke.sk

**Abstrakt** – V tomto článku sú predstavené rôzne metódy biometrickej ochrany s ohľadom na ich výhody a nevýhody a na to, aké sú spôsoby útokov a metód zabezpečenia pri daných útokom. Sú predstavené statické fyziologické aj dynamické behaviorálne znaky. Predstavené sú posledné známe poznatky o najrozšírenejších biometrických bezpečnostných metódach. V závere sú zhrnuté pravdepodobnosti narušenia bezpečnosti rôznych biometrických metód v závislosti od ich typu a sú porovnané proklamované hodnoty miery falošnej akceptácie (FAR) a miery falošného zamietnutia (FRR).

**Kľúčové slová** – biometria, bezpečnosť, FAR, FRR, ERR

## I. ÚVOD

Biometria je veda založená na popise a meraní telesných charakteristík živých bytostí. V súvislosti s automatickými identifikačnými systémami sa biometria vzťahuje na tie systémy a metódy, ktoré sa pri identifikácii alebo autentifikácii spoliehajú na použitie určitej jedinečnej charakteristiky ľudského tela.

Najbežnejšie typy biometrie[1]:

- 1) Statické (fyziologické vlastnosti)
  - Odtlačky prstov
  - Dúhovka oka
  - Sietnica
  - Krvné riečište ruky či prsta
  - Tvár
  - Geometria ruky
  - DNA
- 2) Dynamické - behaviorálne (väčšinou charakteristiky správania)
  - Dynamika písma a podpisu
  - Srdcový rytmus
  - Rytmus hlasu a reči
  - Rozpoznávanie gest
  - Dynamika písania na klávesnici počítača
  - Chôdza

Na meranie spoľahlivosti biometrických systémov sa najbežnejšie používajú FAR (false acceptance rate) miera falošného prijatia, kde  $FAR = \frac{\text{počet falošne prijatých}}{(\text{počet správne zamietnutých} + \text{počet falošne prijatých})}$  teda sleduje hlavne rozhodovanie systému vzhľadom k neznámemu / falošnému používateľovi či útočníkovi (impotor). Druhým je FRR (False rejection rate) miera falošne zamietnutých, teda používateľovi s pravou identitou (genuine) bol prístup odmietnutý,  $FRR = \frac{\text{počet falošne zamietnutých}}{(\text{počet správne prijatých} + \text{počet falošne zamietnutých})}$ . Pri porovnávaní rôznych systémov sa často používa aj miera EER (equal error rate) čo je vybalancovaná hodnota prahu, kde miera chybovosti FAR je zhodná s FRR.

## II. PREHL'AD POUŽÍVANÝCH BIOMETRICKÝCH ZNAKOV A METÓD OCHRANY PRED NAJBEŽNEJŠÍMI ÚTOKMI

### A. Odtlačky prstov

Rozpoznávanie odtlačkov prstov bolo jednou z prvých široko nasadených biometrických metód. V minulosti sa forenzný odbor zaoberajúci sa zberom a analýzou odtlačkov prstov atramentom na

papieri nazýval "daktyloskopia". Prvýkrát bol použitý 18. apríla 1902 v Londýne na identifikáciu zločincov a v priebehu 20. storočia sa používal v kriminalistike v mnohých krajinách. Je založené na identifikácii štruktúry čiar na končekoch prstov, inak známych ako vzory odtlačkov prstov. Po prečítaní skenerom sa jedinečný vzor premení na digitálny biometrický vzor, ktorý systém použije na určenie osoby, ktorá sa nachádza pred ním. Takéto skenery sa delia na dva hlavné typy: optické a kremíkové (tepelné a kapacitné). Každý typ má svoje výhody a nevýhody. Napríklad optické skenery sú najpresnejšie z hľadiska detekcie vzorov, ale dajú sa oklamať použitím silikónových alebo latexových prekrytí a iných trikov. Na rozdiel od lineárnych termálnych senzorov sa rýchlo zašpinia a po každom použití sa musia vyčistiť, aby sa odstránili chyby. Jediný rozdiel pre používateľa je v tom, ako so skenerom komunikuje - dotykom alebo potiahnutím. [2]

1) *Metódy útoku:* Rozdiely medzi falošnými odtlačkami prstov spočívajú vo všeobecnosti v materiáloch použitých na ich vytvorenie. Najčastejšie sa používa technická želatína, hlina, plastelína a zubná sadra. Po získaní vzorky odtlačku prsta používateľa s prístupom do napadnutého biometrického systému sa vytvorí forma, do ktorej sa odleje falošný prst.

2) *Metódy zabezpečenia:* Na určenie toho, či ide o živý odtlačok prsta, sa používajú hardvérové alebo softvérové metódy, prípadne ich kombinácia. Hardvérové metódy:

- Používa sa multispektrálna registrácia (záznam odrazeného IR žiarenia - z kože a zo syntetického materiálu sa získavajú úplne odlišné hodnoty). Zvyčajne sa používa v optických čítačkách;
- záznam impulzov na základe optickej alebo ultrazvukovej metódy;
- meranie elektrického odporu kože.

### B. Krvné riečište na prstoch/rukách

Tento typ je vylepšenou verziou predchádzajúceho prístupu. Jeho algoritmus je oveľa ťažšie prelomiť ako iné biometrické skenovanie, pretože žily sú hlboko pod kožou. Infračervené lúče prechádzajú cez povrch kože, kde sú absorbované žilovou krvou. Špeciálna kamera zachytí obraz, digitalizuje údaje a potom ich buď uloží, alebo použije na potvrdenie totožnosti. Prednedávnom sa táto technológia s názvom FingoPay testovala v jednom londýnskom bare. Systém identifikuje jedinečný vzor žíl na prstoch a podľa výrobcov ide o takmer dokonalé zariadenie, pretože zhoda vzoru žíl u dvoch rôznych ľudí je 1 ku 3,4 miliardy [3].

1) *Metódy útoku:* Biometrická identifikácia cievného riečiška ruky sa považuje za jednu z najodolnejších technológií voči falšovaniu. Ochrana je spôsobená predovšetkým tým, že vzor žily je možné získať len pomocou špeciálneho zobrazovania v infračervenom rozsahu. Na rozdiel od odtlačkov prstov nie je možné zanechať vzory žíl na predmetoch v domácnosti, čo narušiteľovi sťažuje získanie biometrických údajov na falšovanie.

Najzrejmejším spôsobom útoku je získať vzorku na rozpoznanie fotografie cievného lôžka ruky, napríklad v prípade, že je kompromitovaný prostredníctvom ukradnutej databázy.

2) *Metódy zabezpečenia:* Vývojári v prípade kompromitovaného obrazu žily na ruke zabezpečili vyhodnotenie pulzácie krvného toku v žilách, čo veľmi sťažuje vytvorenie modelu ruky. Algoritmus na detekciu falšovania je viacstuňový. V prvej fáze sa určí, či je prezentovaný objekt živý, konkrétne sa analyzujú koeficienty odrazu/absorpcie IR svetla v dvoch spektrálnych rozsahoch materiálom objektu. V ďalšej fáze sa analyzuje tvar objektu, aby sa určilo, že ide o dlaň/prst, a nie o inú časť ľudského tela.

### C. Geometria dlane

Určenie geometrie ruky sa týka merania charakteristík, ako je dĺžka a šírka prstov, ich zakrivenie a relatívna poloha. V súčasnosti je táto metóda zastaraná a už sa takmer nepoužíva, hoci kedysi bola dominantnou jednoducho dostupnou verziou biometrickej identifikácie (letisko Košice). Moderný pokrok v oblasti softvéru na rozpoznávanie odtlačkov prstov a tváre zatienil jeho význam. Táto metóda sa však dodnes používa veľmi zriedkavo, a to z dôvodu jej nepresnosti a dostupnosti modernejších biometrických technológií [2].

### D. Dúhovka

Dúhovka, teda farebná časť oka, sa skladá z hrubých vláknitých svalov. Tieto svaly pomáhajú formovať zrenicu a kontrolovať množstvo svetla, ktoré vstupuje do oka. Meraním jedinečných záhybov a charakteristík týchto svalov môžu biometrické overovacie nástroje potvrdiť totožnosť s neuveriteľnou presnosťou. Technológie dynamického skenovania (napríklad skenovanie žmurkania) zvyšujú úroveň presnosti a bezpečnosti [4].

1) *Metódy útoku:* Sú známe dve metódy: ukázať skeneru fotografiu dúhovky alebo vyrobiť kontaktné šošovky s imitáciou dúhovky.

2) *Metódy zabezpečenia:* V súčasnosti konštruktéri moderných biometrických systémov používajú na získanie tejto biometrickej charakteristiky ožarovanie očí v blízkom infračervenom pásme. Získanie tejto biometrickej charakteristiky v ľubovoľných podmienkach bez použitia špeciálneho zariadenia je však pomerne problematické. Na ochranu sa používa aj pozorovanie ľubovoľných a vynútených zmien dúhovky oka (reakcia na svetlo).

#### E. Sietnica

Skenovanie sietnice umožňuje skenovať kapiláry hlboko vo vnútri oka pomocou kamier s blízky infračerveným žiarením. Výsledný obraz sa najprv predbežne spracuje, aby sa zlepšila jeho kvalita, a potom sa prevedie na biometrickú šablónu na registráciu nového používateľa a následné porovnanie referencií počas pokusov o rozpoznanie používateľa [5]. Širšiemu používaniu takýchto skenerov bránia vysoké náklady a potreba umiestniť oko blízko kamery [4].

#### F. Tvár

Technológia rozpoznávania tváre je určite jednou z najmodernejších foriem biometrických identifikačných systémov. Tento druh softvéru meria geometriu tváre vrátane vzdialenosti medzi očami a od brady po čelo (a to je len niekoľko parametrov). Po zozbieraní údajov ich pokročilý algoritmus prevedie na zašifrovaný vzor, inak známy aj ako podpis tváre (signatúra). Podľa správy spoločnosti Computer Sciences Corporation (CSC) už mnohé obchody zaviedli systémy na rozpoznávanie tváre na sledovanie určitých skupín zákazníkov. Princíp fungovania tohto typu systémov možno prirovnať k cieľenej reklame, ktorej cieľom je skúmať preferencie a ponúkať najrelevantnejšie produkty. V poslednom čase sa táto technológia stala veľmi populárnou medzi používateľmi smartfónov vďaka rôznym aplikáciám (napríklad na určenie veku) alebo zabudovaným skenerom tváre, ktoré umožňujú odomknutie zariadenia. Vzhľadom na rozšírenosť medzi používateľmi a relatívnu jednoduchosť používania tejto metódy sa začalo objavovať čoraz viac aplikácií využívajúcich túto technológiu. V roku 2019 došlo k niekoľkým škandálom v dôsledku úniku údajov z aplikácie Zao na falošné fotky (deepfake). V reakcii na to 16. marca 2020 vyšlo najavo, že Čína zaviedla aktualizované normy pre aplikácie, ktoré zhromažďujú biometrické údaje vrátane systémov na rozpoznávanie tváre [6].

1) *Metódy útoku:* Najznámejšie metódy falšovania biometrie tváre sú:

- zobrazenie fotografie tváre;
- prezentácia nahraného videozáznamu tváre;
- používanie masky;
- aplikácia make-upu;
- výroba formy (modelu) hlavy človeka;
- prezentácia počítačom vytvoreného videomodelu tváre osoby.

2) *Metódy zabezpečenia:* Na boj proti falšovaniu sa používajú aj hardvérové a softvérové metódy a ich kombinácie. Pulz sa môže merať na videozázname tváre a sledovanie zmien obrazu, keď sa požiada o otočenie hlavy alebo žmurkanie.

#### G. Tvar zvukovodu ucha

Na rozdiel od mnohých iných biometrických metód, ktoré si vyžadujú špeciálne kamery, tieto biometrické systémy merajú akustiku ucha pomocou špeciálnych slúchadiel a nepočuteľných zvukových vln. Mikrofón v každom slúchadle meria, ako sa zvukové vlny odrážajú od ušnice a rozchádzajú sa rôznymi smermi v závislosti od zakrivenia zvukovodu. Digitálna kópia tvaru zvukovodu ucha sa prevedie na biometrickú šablónu na neskoršie použitie. V súčasnosti ešte nie sú k dispozícii komerčné senzory, ale v tejto oblasti prebieha intenzívny výskum. Na jar 2015 spoločnosť Yahoo Labs navrhla identifikovať majiteľa smartfónu na základe jeho tvaru zvukovodu ucha. V tom istom roku získala patent na podobný prístroj aj spoločnosť Amazon [7].

#### H. Hlas

Technológia rozpoznávania hlasu patrí do oblasti fyziologickej aj behaviorálnej biometrie. Z fyziologického hľadiska takéto systémy rozpoznávajú tvar ľudského hlasového ústrojenstva vrátane nosa, úst a hrtana a určujú vydávaný zvuk. Z behaviorálneho hľadiska zaznamenávajú, ako človek niečo hovorí - zmeny v pohybe, tóne, tempe, prízvuku atď., čo je tiež jedinečné pre každého jednotlivca. Kombináciou údajov z fyzickej a behaviorálnej biometrie sa vytvorí presný hlasový podpis, hoci sa môžu vyskytnúť určité nezrovnalosti (napr. v prípade choroby alebo iných faktorov) [8].

1) *Metódy útoku*: Najjednoduchší spôsob, ako podvrhnúť hlas, je použiť diktafón s hlasom požadovanej osoby. Moderné biometrické systémy vás však pravdepodobne vyzvú, aby ste povedali nejaké neznáme heslo, alebo vás požiadajú, aby ste povedali nejakú neznámú frázu. V tomto prípade by ste museli na útok použiť syntetizátor reči. Rozlišujú sa tieto typy útokov na systémy rozpoznávania hlasu:

- napodobňovanie hlasu (imitácia);
- nahrávanie hlasových biometrických charakteristík osoby a ich následné opakovanie;
- premena reči útočníka na reč inej osoby;
- syntéza reči.

2) *Metódy zabezpečenia*: Prvé dve metódy sa dajú ľahko odhaliť tak, že požiadate osobu, aby povedala prístupovú frázu. Najväčšou hrozbou pre hlasové biometrické systémy je technika konverzie hlasu, teda prevod reči útočníka na reč klienta biometrického systému, a útok syntézou reči, ktorý spočíva vo vytvorení syntetického hlasu používateľa a verifikačného systému.

### I. Termogram tváre

Termogram je zobrazenie infračervenej energie tváre vo forme obrazu rozloženia teploty. Biometrická termografia tváre zachytáva tepelné vzory spôsobené pohybom krvi pod kožou. Keďže cievy každého človeka sú jedinečné, aj príslušné termogramy sú jedinečné, dokonca aj medzi identickými dvojčatami, vďaka čomu je táto biometrická metóda overovania ešte presnejšia ako tradičné rozpoznávanie tváre [6].

### J. DNA

DNA sa už dlho používa ako metóda identifikácie. Okrem toho je to jediná forma biometrie, ktorá dokáže sledovať rodinné väzby. Porovnávanie DNA je obzvlášť cenné pri riešení prípadov nezvestných osôb, identifikácii obetí katastrof a potenciálneho obchodovania s ľuďmi. DNA získaná z vlasov, slín a podobne, obsahuje sekvencie krátkych tandemových opakovaní (STR). Možno ich použiť na jednoznačné potvrdenie identity porovnaním s inými STR v databáze. Táto technológia je v súčasnosti na trhu s biometriou málo zastúpená. DNA sa považuje za ideálnu biometrickú charakteristiku, ale jej nedostatkom je, že identické dvojčatá majú rovnakú DNA [2].

## III. BEHAVIORÁLNE SYSTÉMY

Princípy behaviorálnej biometrie sú založené na charakteristikách ľudského pohybu a správania. Nižšie sú uvedené hlavné typy behaviorálnej biometrie [9].

### A. Vlastnoručný podpis

Rozpoznávanie podpisu je behaviorálny biometrický systém, ktorý meria priestorové súradnice, tlak pera, náklon pera a ťah pera v samostatných aj interaktívnych aplikáciách. Digitálny tablet zaznamenáva merania a potom tieto informácie využíva pri automatickom vytváraní biometrického profilu na budúce overovanie. V súčasnosti sa na zadávanie podpisu používajú tablety, ktoré automaticky zaznamenávajú polohu pera v rôznych časových bodoch, uhly sklonu a tlak vyvíjaný na tablet [9].

### B. Dynamika stlačenia klávesov

Dynamika stlačenia klávesov posúva štandardné heslá na novú úroveň vďaka sledovaniu rytmu ich zadávania. Takéto snímače môžu reagovať na čas potrebný na stlačenie každého klávesu, oneskorenie medzi klávesmi, počet znakov zadaných za minútu atď. Vzory stlačenia klávesov sa používajú v kombinácii s heslami a kódmi PIN na zlepšenie zabezpečenia [9]. Využíva sa aj audio monitoring [10], [11] klávesnice pre multimodálne systémy [12].

## IV. ZÁVER A ZHODNOTENIE

Pravdepodobnosť falšovania nezávisí od typu biometrického prvku, ale od technológie použitej na jeho čítanie. [6] Hodnoty miery falošnej akceptácie (FAR) a miery falošného zamietnutia (FRR) patria medzi najdôležitejšie parametre biometrických systémov a výrobca ich spravidla uvádza medzi kľúčovými vlastnosťami biometrického zariadenia. [6]

Biometrické systémy majú veľmi sľubnú budúcnosť. Aplikácie biometrickej autentifikácie identity zahŕňajú systémy kontroly prístupu, systémy kontroly dodržiavania časového rozvrhu, zabezpečenie pracovnej stanice, prístupu k sieťam a doménam, aplikácie na prihlasovanie, ochranu údajov, vzdialený prístup k zdrojom, podporu vzdialených transakcií a bezpečnosť webu a v poslednom čase sa vo veľkej miere využívajú pri overovaní biometrických podpisov. Už mnoho rokov sa biometria používa v mnohých malých i veľkých spoločnostiach, kanceláriách a priemyselných podnikoch. Investície do nákupu biometrických čítačiek výrazne stúpajú a používanie biometrických údajov sa stalo bežným.



Tabuľka 1  
Pravdepodobnosť manipulácie pri rôznych biometrických technológiách.

Biometrická čítačka	Pravdepodobnosť manipulácie
Multifunkčná čítačka odtlačkov prstov	nízka
Optická čítačka odtlačkov prstov	možná
3D čítačka odtlačkov prstov	nízka
Detekcia tváre 2D	možná
Detekcia tváre 3D	možná
Krvné riečište	nízka
Sietnica	nízka

Tabuľka 2  
Priemerná chybovosť pri najbežnejších biometrických systémoch.

Biometrická metóda verifikácie	FAR	FRR
Odtlačok prsta	0,001%	0,6%
Dúhovka	0,00001%	0,016%
Detekcia tváre 2D	0,1%	2,5%
Detekcia tváre 3D	0,0005%	0,1%
Krvné riečište	0,0008%	0,01%
Sietnica	0,0001%	0,4%

#### ACKNOWLEDGMENT

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja vďaka projektu APVV SK-TW-2017-0005, Kultúrnou a edukačnou grantovou agentúrou Ministerstva školstva vedy výskumu a športu Slovenskej republiky vďaka projektu číslo KEGA 009TUKE-4/2019 s názvom „Inovácia obsahu a príprava učebných textov pre predmet Biometrické systémy bezpečnosti“, Vedeckou grantovou agentúrou realizáciou výskumného projektu VEGA 1/0753/20 „Robustné rečové technológie metódami hlbokého učenia“ a VEGA 2/0165/21 „Technológie automatického spracovania reči na pomoc v krízových situáciách“ financovaných z prostriedkov Ministerstva školstva, vedy, výskumu a športu SR, Asociáciou COST v projekte CA16101 MULTI-modal Imaging of FOREnsic SciEnce Evidence (MULTI-FORESEE) - tools for Forensic Science a nakoniec aj cez Operačný program Výskum a inovácie, financovaného z Európskeho fondu regionálneho rozvoja v rámci projektov „Univerzitný vedecký park TECHNICOM pre inovačné aplikácie s podporou znalostných technológií – II. fáza“ Kód ITMS: 313011D232 a „Inteligentné dátové infraštruktúry pre digitálnu spoločnosť“ kód ITMS: 313011T571.

#### LITERATÚRA

- [1] A. Babich, “Biometric authentication. types of biometric identifiers,” *Bachelor’s Thesis*, 2012, HAAGA-HELIA.
- [2] A. C. Weaver, “Biometric authentication,” *Computer*, vol. 39, no. 2, pp. 96–97, 2006.
- [3] I. Sarkar, F. Alisherov, T.-h. Kim, and D. Bhattacharyya, “Palm Vein Authentication System: A Review,” *International Journal of Control and Automation*, vol. 3, no. 1, pp. 27–34, 2010.
- [4] M. Douglas, K. Bailey, M. Leeney, and K. Curran, “An overview of steganography techniques applied to the protection of biometric data,” *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 17 333–17 373, 2018.
- [5] M. Messerschmidt and M. Pleva, “Advanced biometric systems utilizing neural networks in identity evaluation,” in *2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA)*. IEEE, 2020, pp. 458–463.
- [6] D. Bhattacharyya, R. Ranjan, F. Alisherov, M. Choi *et al.*, “Biometric authentication: A review,” *International Journal of u-and e-Service, Science and Technology*, vol. 2, no. 3, pp. 13–28, 2009.
- [7] I. M. Alsaadi, “Physiological biometric authentication systems, advantages, disadvantages and future development: A review,” *International Journal of Scientific & Technology Research*, vol. 4, no. 12, pp. 285–289, 2015.
- [8] A. Kounoudes, V. Kekatos, and S. Mavromoustakos, “Voice biometric authentication for enhancing internet service security,” in *2006 2nd International Conference on Information & Communication Technologies*, vol. 1. IEEE, 2006, pp. 1020–1025.
- [9] Z. Jorgensen and T. Yu, “On mouse dynamics as a behavioral biometric for authentication,” in *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, 2011, pp. 476–482.
- [10] E. Vozáriková, M. Pleva, J. Juhár, and A. Cizmár, “Surveillance system based on the acoustic events detection,” *Journal of Electrical and Electronics Engineering*, vol. 4, no. 1, p. 255, 2011.
- [11] M. Pleva, E. Kiktova, J. Juhar, and P. Bours, “Acoustical user identification based on mfcc analysis of keystrokes,” *Advances in Electrical and Electronic Engineering*, vol. 13, no. 4, pp. 309–313, 2015.
- [12] M. Pleva, P. Bours, S. Ondáš, and J. Juhár, “Improving static audio keystroke analysis by score fusion of acoustic and timing data,” *Multimedia Tools and Applications*, vol. 76, no. 24, pp. 25 749–25 766, 2017.

# Audiovisual Live Coding with Audience Participation via Parameter Guessing

<sup>1</sup>Daniel MAJERNÍK, <sup>2</sup>Matúš SULÍR

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovakia  
<sup>1</sup>daniel.majernik@student.tuke.sk, <sup>2</sup>matus.sulir@tuke.sk

**Abstract** – Live coding is a process of writing and executing source code in real-time to generate an artistic performance. In addition to passive observation, a few live coding systems allow the audience to actively participate in the performance in some ways. This paper introduces a novel approach to audience participation in live coding environments – in the form of a parameter guessing game. In the proposed system, a performer is live coding to create audio and visual elements of a performance. The participants are assigned random parameters of the elements (e.g., a box color or cymbal volume). They can modify these parameters without knowing what the actual parameter is. Their goal is to correctly guess which parameter they are modifying. We designed and implemented a web application encompassing the mentioned principles. The structure of a performance is created by a performer using a simple live coding language we devised. Based on the performed user evaluation, we conclude that the system is exciting for the participants, however, the performer’s interface needs to be improved.

**Keywords** – audience participation, audiovisual performance, guessing game, live coding

## I. INTRODUCTION

Live coding is a unique form of artistic expression. Performers in live coding performances write code, which generates audio, visual, or audiovisual pieces. For live coding performances, a manifesto has been published to define what principles such a performance should follow [1]. Many, if not all the live coding environments have been built upon these principles. One such environment is LiveCodeLab [2]. It provides realtime audiovisual performances, along with a very interesting feature called *autocode*, which automatically generates a visual performance. Gibber is a complex audiovisual live coding environment, which uses JavaScript for live coding as a way of learning a real programming language along with the actual live coding [3]. It is available in a web browser, but it is not suited for mobile usage. Another great example is Sonic Pi, which uses the Ruby language and has extensive musical capabilities. It has a remarkable sound quality and its syntax is relatively simple, as it was intended to be used for education purposes [4]. Lan and Jensenius developed a web-based system for collaborative live coding of audio performances - QuaverSeries [5]. Their idea behind a web-based solution was that it is far more accessible than software needing installation. There are many examples of amazing performances created using the mentioned systems available on YouTube.

Audience participation in live coding introduces a completely new, outer factor to performances. Research by Turchet et al. brings up important considerations, the key points for these systems being synchronicity and low latency [6]. The use of mobile devices as participation intermediaries was successfully implemented by Lee and Willette. In their system, the participants created their melodies and subsequently browsed other participants’ melodies as a metaphor for online dating [7]. The performer’s input was only in a form of changing the overall scale in which the melodies were playing. Another interesting system is OpenBand, where users create sounds from typing messages. It uses its own audio typography to synthesize various sounds from different messages [8]. Both these systems used Web Audio API. A different form of approach is using gestures on mobile devices. This was the subject of Handwaving, which used neural networks to process data from accelerometers [9]. It was applied to audio performances in web browsers and they have not had any trouble running them on mobile devices with a production date of at least 2017 [9]. Different, but in their nature similar systems, are collaborative drawing boards. Such a system was devised by Adianto et al. for education purposes [10]. By using the *Websocket* protocol, they

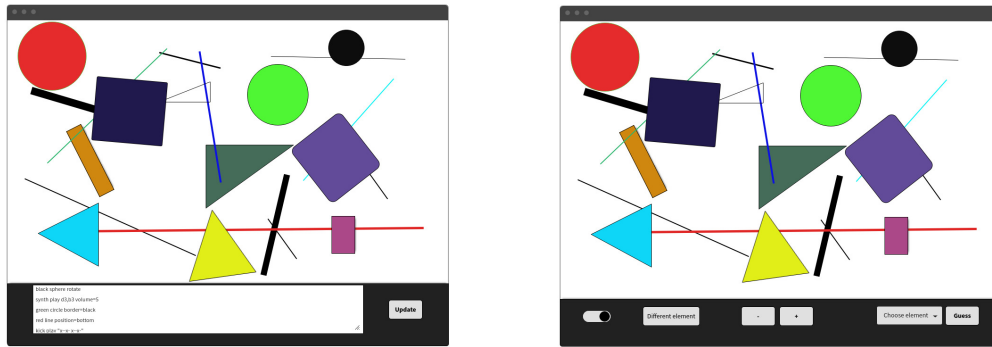


Fig. 1 Prototype of the performer's (left) and participant's (right) user interface

achieved a latency of less than 2.2 ms, which easily meets the requirements found in the research of Turchet et al. [6].

The goal of this paper is to describe the development of a system in which a performer is live coding an audiovisual performance. The participants can modify certain elements of it without reducing the performance to incomprehensible chaos. Besides the modification of the performance itself, the system contains a game component, as these participants do not know which parameter they are modifying, and their aim is to guess it. We also discuss the ideas behind certain choices in the development process. Finally, we review the user testing, conclude the results, and draw some ideas for further research.

## II. PROTOTYPING THE APPLICATION

To prototype the application, we drew from the functional requirements. Because of the emerging pandemic, the goal was also to be able to broadcast the performance for the possibility of performing online. Therefore, the idea of a single-page application came to be. The audio aspect is irrelevant as to designing the interface. To ensure the visibility of all visual elements, the scene on which they are being rendered takes up the full width of a web page and approximately 70-80% of its height. This is the same for participants and performers, so there are no inconsistencies in the displayed performance. The difference is the bottom part, where performers need to have their live coding input, as seen in the left part of Fig. 1. In the case of participants, there are located all the elements necessary to interact with the performance, as displayed in the right part of Fig. 1.

## III. DESIGN OF THE LIVE CODING LANGUAGE

In an audiovisual performance, there are 2 different kinds of elements. This notion led to discerning the commands by their two respective elements: audio and visual commands. Audio commands are preceded by the word *play*, whereas visual commands start with *draw*. Individual commands are separated by an empty line, which puts them into the form of blocks or paragraphs. This allows the live coder to style commands in any way most comprehensive or expressive to the performer himself. Strings within a command are separated by empty characters, e.g., spaces.

Audio commands need to have an instrument to play the part. To that end, the second string of a command is always the preferred instrument's name. Additionally, there is a possibility of declaring a custom name. This name is defined by concatenating a character ':' and specifying the custom name, e.g., *synth:customName*. The subsequent word declares the play type, i.e. if single notes or chords should be played. The trickier section is creating the musical score or rather the notes and their durations to be played. This constitutes the fourth string of an audio command and the score is generated by using the following characters:

- *note* – defining the note to be played,
- *character* '-' – is used to prolong the duration of a preceding note,
- *character* 'x' – stands for silence in the part, acts as a musical rest.

Note names are represented as their musical notation, e.g. a4, Eb3, c#2, where the number means the octave in which they are to be played. The letter case is irrelevant for playing single notes, although there is an exception when declaring chord names. A common approach is to use uppercase letters for major chords, and lowercase letters for minor chords, which is what we implemented in this language. All the specified characters are inside '|' characters, which

instinctively represent musical tacts. There is also an empty tacts notation which is useful to play parts in different time of the musical piece. It is achieved by typing a digit in the tact, defining the number of empty tacts. The musical part representation is inspired by LiveCodeLang of LiveCodeLab environment [2]. However, in LiveCodeLang, there are no prolonged durations of notes and no different note definitions. An example of a correct audio command of the designed language is depicted in Listing 1.

Listing 1 Audio command example in the proposed language

```
play polysynth:my_chords chords |4|a3-|4|d3 d3 d3 d3|
```

Visual commands have a fairly simple structure, as they do not have to follow any kind of rules music requires. The approach to generate them is pretty straightforward. The first word after *draw* is a shape of the element. The shape can be either a geometric shape or a light. Subsequently, the rest of the command is typed in paired strings of the following format: *parameter:value*. This allows for consistency and simplicity of potential expansion of new parameters. An example of a correct audio command of the designed language is depicted in Listing 2.

Listing 2 Visual command example in the proposed language

```
draw box position:1,2,-10 size:5,5,5 color:blue rotate:0.01,0.01,0  
name:box_in_background
```

#### IV. IMPLEMENTATION

The developed application works on the principle of rooms. Upon arriving at the home page, a user can join a room. In each room, there is a different performance running. If he joins a non-existent room, a room is created and he becomes the performer in it. Conversely, if he joins an existing room, he becomes a participant in its performance.

The communication between all participants of the performance is managed by the *Node.js* runtime, which is capable of running tens of thousands of concurrent connections [11]. It also makes the whole implementation concise using only JavaScript. Data is transferred over the *WebSocket* protocol. To establish this communication, we utilized the *socket.io* library. It provides a simple way to create rooms for multiple performances running simultaneously.

To implement the user interface on the web, we chose the *React.js* library. It has been successfully used to develop collaborative live coding environments [5][12], which share a certain resemblance to this system. Because of the nature of the prototyped interface, we opted for a single-page application. This is a built-in feature of *React.js*. The next reason behind this library is the conditional rendering it provides. It allows for a different interface within the same component based on whether a participant or a performer is using it. To generate sounds and visuals, we programmed a script to act as a mediator between the interface and the libraries responsible for generating them. The interfaces of performers and participants can be seen respectively in Fig. 2 and Fig. 3.

Audio elements are synthesized using the *Tone.js* framework, which is built on Web Audio and provides DAW-like<sup>1</sup> features for coding music in web browsers. We support all of the instruments it provides. To render visual elements, we used the popular *Three.js* library. We utilized most of the geometric shapes and lights it provides. Both of the libraries are well documented and accessible by all of the most popular web browsers, even on mobile devices. Lee et al. used *Three.js* in a real-time modification of a poem being typed [13].

To ensure the circulation of parameters we implemented a FIFO (first in, first out) queue to hold the unassigned parameters, and another FIFO queue to hold the guessing users. The idea behind this is simple: we wanted to ensure the users being the longest without a parameter to receive one the first. Every participant, who starts guessing, is added to the end of the queue. If some participant changes his parameter, it is immediately assigned to the first user in the queue. If there is no user in the queue, the parameter is added to the queue of unassigned parameters. These are distributed every time a user starts guessing. When the performer adds elements to the performance, its new modifiable parameters are distributed, whereas the leftover is added to the queue. Also, the user who changes his parameter does not get the same one if there is no other parameter available. This way, if someone else changes his parameter, both of them get a different one. This sort of switching parameters between each other ensures a concise and quick distribution without the need of setting up any kind of interval-based function calls. The one border case with this approach is users disconnecting whilst being assigned a parameter.

<sup>1</sup>DAW - digital audio workstation

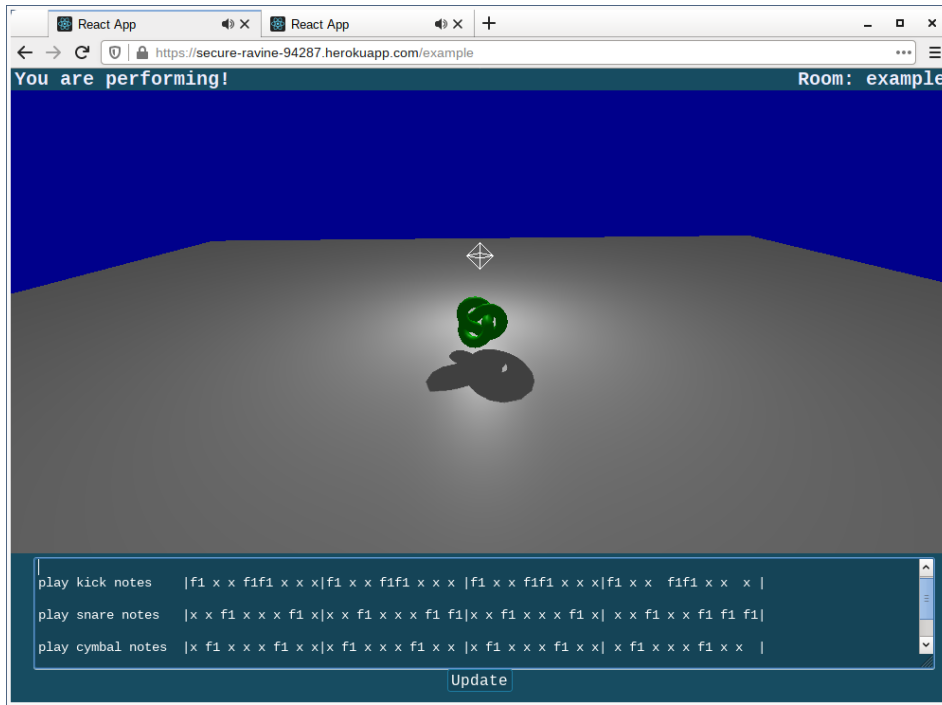


Fig. 2 Performer's user interface

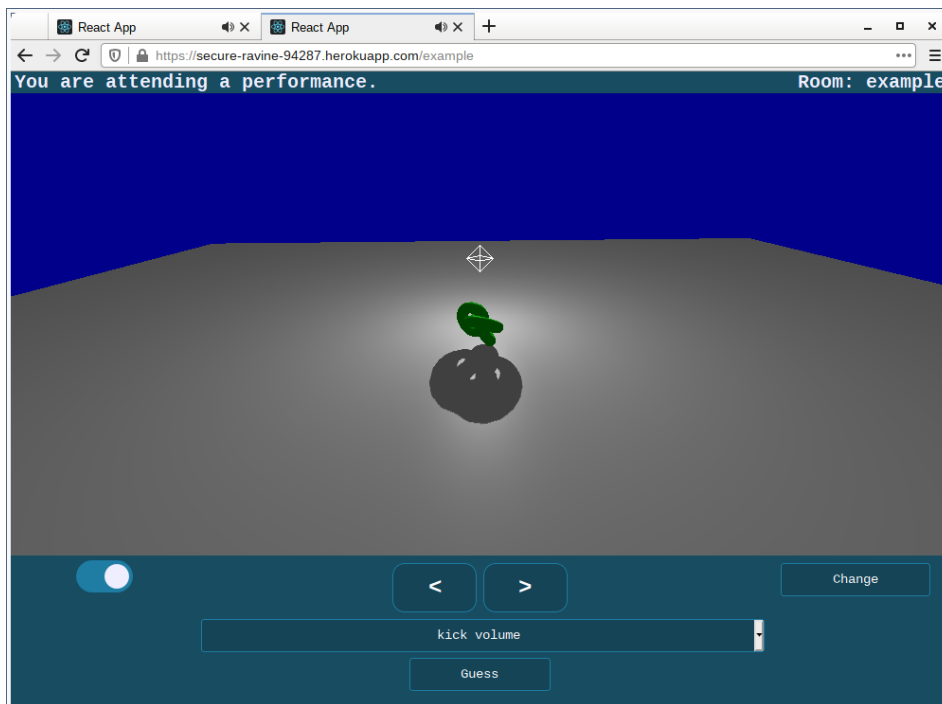


Fig. 3 Participant's user interface

This would lead to removing the parameter from circulation. Therefore, we have to save the IDs of the users and their respective parameters on the server to avoid losing the parameter. When a performer removes certain elements of a performance, the parameters belonging to them are removed from the queue on the server and also deleted locally, if a participant happens to have it assigned. This way the server does not have to retrieve assigned parameters and delete them separately.

To serve the client files and set up the routing to different components we used the *Express* framework. Under the hood, it uses the standard HTTP module, which is subsequently upgraded to WebSocket connection via the *socket.io* library. The WebSocket protocol achieves the best results in most performance measurements of data transmission over the internet [14][15][16],

which allows for remarkably low latency. The socket.io usage can be observed in Soundworks – an audiovisual performance which uses gesture on mobile devices [17]. This library is attached to the Express instance and handles all the events pertaining to the performance.

The application has been deployed to the Heroku platform. You can find the current version of the application at <https://secure-ravine-94287.herokuapp.com>.

## V. USER TESTING

Two tests were performed to evaluate the usability and user experience of the developed system. The results of each test were acquired in two ways: a discussion with the users immediately after the test and presenting them with a questionnaire based on SUS (system usability scale). Afterward, we draw conclusions from the results.

### A. Audience

The first test was focused on the participants (audience). It was performed on a local WiFi network with 4 participants and the first author as a live coder. Live coding was performed on a laptop using Microsoft Edge as the web browser of choice. The laptop was connected to the LG 50PK350 television, which served as the audio and visual source of the performance. The participants were connected using Google Chrome web browser on their mobile devices, each with a production year of at least 2019. None of the participants knew or even heard about live coding before. Therefore, we had briefly informed them about it and what is the nature of their involvement before the performance commenced.

At the start of the performance, the participants were more interested in what the performer was doing. The guessing was not very interesting to them as there were not many elements in the performance and they guessed right really fast. Users who could not speak English were asking about the selection of guesses they were provided, mostly from the beginning. However, as soon as there were more elements than the number of users, the commotion started. They almost completely lost interest in the live coding and focused mainly on guessing their parameters. Strong competition between the users was observed during the second half of the performance. A certain level of frustration was noticeable when an element whose parameter was assigned to some user was removed. After that, the performer tried to predominantly add elements and not remove them. The performance ended after approximately 20 minutes.

The feedback from the discussion was mostly positive. Two of the participants would like to engage in a live coding performance again. They collectively agreed that a notification should be present when their parameter is removed. Two of the users would also like to have a translation for parameter types. One participant complained about rotation modification being unnoticeable. The SUS score of the questionnaire was 81.25. This means above-average usability. There was only one case of an under-average score, 65, narrowly missing the borderline score of 68.

### B. Performer

To test the performer's interface, one user tried to live code within the application. He had no previous experience with live coding; however, he had experience with regular coding. During his test, there was no audience present. His web browser of choice was Google Chrome on his personal computer, and the test lasted approximately 30 minutes.

During the discussion, the tester was enthusiastic about adding new elements to the performance, but he was not satisfied with the overall interface. He pinpointed the text field for live coding commands as the most disliked feature. The main point of his interest was the live coding of the audio elements. He tried to create a melody of a song he knew with a fairly satisfactory result. The SUS score of the questionnaire was 52.5. This value means moderately severe issues with the interface, which requires further investigation to improve its usability.

## VI. CONCLUSION

The user testing, although not very extensive, brought up some valid points. The first one is the difference in the evaluation of the participants' and performer's user interface. A minor upgrade to participants' user interface to alleviate the frustration of removed parameters is necessary. As for the live coding interface, we would propose a more advanced text editor, to begin with. The idea was that a simple live coding language would not require a full-featured source code editor. However, the testing proved this assumption to be incorrect. Nonetheless, it would definitely be of great relevance to gather more data to work with – testing by skilled live coders, with a greater number of participants, online performances with users from around the globe, etc. Furthermore,

the designed live coding language is currently simple, so it could be enriched with programming constructs such as loops or conditionals.

For the participants, the performance balanced on the verge of a game. Is that a success or a failure in terms of a live coding performance? One could argue that if the participants enjoyed their time at any performance, it was a success. During the guessing and modifying, they have to be absolutely focused on every element of the performance. From the viewpoint of a performer it might seem a bit different, but allow us to make an analogy to the more traditional art forms: You do not need to see the guitar of a musician, nor do you need to see the brush of a painter, to enjoy their art. Ultimately, it is up to individual participants, whether they want to look at them during the performance. It would be a good point for further research to immerse participants more deeply in the live coding aspect of the performance. We could summarize the performances created within this system as a result of a fascinating collaboration of the participants playing a game, and the performer building a playground for them.

## REFERENCES

- [1] "TOPLAP manifesto." [Online]. Available: <http://toplap.org/wiki/ManifestoDraft>
- [2] D. Della Casa and G. John, "LiveCodeLab 2.0 and its language LiveCodeLang," in *Proceedings of the 2nd ACM SIGPLAN International Workshop on Functional Art, Music, Modeling & Design*, ser. FARM '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 1–8.
- [3] C. Roberts, M. Wright, J. Kuchera-Morin, and T. Höllerer, "Gibber: Abstractions for creative multimedia programming," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 67–76.
- [4] S. Aaron, "Sonic Pi – performance in education, technology and art," *International Journal of Performance Arts and Digital Media*, vol. 12, no. 2, pp. 171–178, 2016.
- [5] Q. Lan and A. R. Jensenius, "QuaverSeries: A live coding environment for music performance using web technologies," in *Proceedings of the International Web Audio Conference (WAC)*. NTNU, 2019, pp. 41–46.
- [6] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of musical things: Vision and challenges," *IEEE Access*, vol. 6, pp. 61 994–62 017, 2018.
- [7] S. W. Lee and A. Willette, "Crowd in C," in *Proceedings of the 2019 on Creativity and Cognition*, ser. C&C '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 425–431.
- [8] A. Stolfi, M. Barthet, F. Goródsy, A. Deusany, and F. Iazzetta, "Open band: Audience creative participation using Web Audio synthesis," in *Proceedings of the International Web Audio Conference*, ser. WAC '17, F. Thalmann and S. Ewert, Eds. Queen Mary University of London, 2017.
- [9] G. Roma, A. Xambó, and J. Freeman, "Handwaving: Gesture recognition for participatory mobile music," in *Proceedings of the 12th International Audio Mostly Conference*. Association for Computing Machinery (ACM), 2017.
- [10] M. F. Adianto, M. Ilmam, A. K. Miharja, Y. Bandung, and L. B. Subekti, "Design and implementation of Cakra Wahana Paksa, a collaborative whiteboard application based on WebSocket technology for distance learning," in *2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2016, pp. 1–6.
- [11] J.-r. Wang, C.-p. Xu, and X.-q. Xi, "Design and implementation of a course selection system that based on Node.js," in *Proceedings of the Fourth Euro-China Conference on Intelligent Data Analysis and Applications*. Cham: Springer International Publishing, 2018, pp. 167–176.
- [12] J. Fiala, M. Yee-King, and M. Grierson, "Collaborative coding interfaces on the Web," in *International Conference on Live Interfaces*, 2016, pp. 49–57.
- [13] S. W. Lee, G. Essl, and M. Martínez, "Live writing : Writing as a real-time audiovisual performance," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2016, pp. 212–217.
- [14] R. Appelqvist and O. Örnmyr, "Performance comparison of XHR polling, long polling, server sent events and Websockets," Bachelor's thesis, 2017. [Online]. Available: <https://www.diva-portal.org/smash/get/diva2:1133465/FULLTEXT01.pdf>
- [15] K. E. Ogundeyi and C. Yinka-Banjo, "WebSocket in real time application," *Nigerian Journal of Technology*, vol. 38, no. 4, pp. 1010–1020, 2019.
- [16] W. Słodziak and Z. Nowak, "Performance analysis of web systems based on XMLHttpRequest, server-sent events and WebSocket," in *Information Systems Architecture and Technology: Proceedings of 36th International Conference on Information Systems Architecture and Technology – ISAT 2015 – Part II*. Cham: Springer International Publishing, 2016, pp. 71–83.
- [17] S. Robaszekiewicz and N. Schnell, "Soundworks—a playground for artists and developers to create collaborative mobile web performances," in *WAC-1st Web Audio Conference*, 2015.

# Riadenie zásobníka energie elektrického midibusu

<sup>1</sup>Dávid BODNÁR, <sup>2</sup>František ĎUROVSKÝ

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

<sup>1</sup>david.bodnar@student.tuke.sk, <sup>2</sup>frantisek.durovsky@tuke.sk

**Abstrakt** — Článok sa zaoberá návrhom a riadením zásobníka energie pre elektrický midibus. Zásobník energie obsahuje batériu a superkapacitor. Úlohou superkapacitora je znížiť záťaž vyvíjanú na batériu počas akcelerácie a rekuperačného brzdzenia a predĺžiť tak jej životnosť. Bol zostavený systém manažmentu rozdeľovania energie medzi batériu, superkapacitor a motor vozidla, ktorý vychádza z kinetickej a potenciálnej energie vozidla. Model batérie obsahuje aj model starnutia. Funkčnosť zásobníka energie bola overená simuláciou v programe MATLAB. Simulačné výsledky sú podrobne rozanalyzované, vrátane zhodnotenia vplyvu superkapacitora na životnosť batérie, vnútorné straty zásobníka a spotrebu energie vozidla. Práca obsahuje aj metodiku optimalizácie veľkosti komponentov zásobníka energie pre daný jazdný cyklus midibusu.

**Kľúčové slová** — model batérie, superkapacitor, QSS toolbox, manažment energie, zásobník energie, starnutie batérie

## I. ÚVOD

Snaha Európskej únie o dekarbonizáciu hospodárstva sa prejavuje v doprave tlakom na masívne nasadzovanie elektromobilov. Kritickým komponentom elektrických vozidiel (EV) sú stále trakčné batérie. Súčasná batérová technológia neposkytuje takú energetickú hustotu, bezpečnosť a životnosť akumulátorov, aby mohli elektromobily plne konkurovať vozidlám so spaľovacími motormi. Elektromobily sú v súčasnosti vhodné na prevádzku v mestách, kde sa od nich nevyžaduje dlhý dojazd, a s výhodou dokážu vďaka schopnosti rekuperácie využiť kinetickú energiu vozidla pri brzdení. Nočné nabíjanie je šetrné k batériám a zároveň pomáha vyrovnávať odber elektrickej energie z verejnej siete. Zaujímavou možnosťou je aj ich schopnosť akumulovať (dodávať) časť elektrickej energie z verejnej siete, a tak pomáhať vyrovnávať kolísanie dodávky a odberu energie vo verejnej distribučnej sieti. Napriek uvedeným možnostiam si zatiaľ elektrické vozidlá nenašli dostatok zákazníkov medzi bežnými užívateľmi, ktorí dávajú prednosť vozidlám schopným dlhodobo jazdiť aj mimo miest.

Iná situácia je v oblasti mestskej hromadnej dopravy (MHD). Vozidlá MHD jazdia po stálych trasách s častými rozjazdmi a brzdeniami. Navyše jazdia v oblasti, kde sa kladie dôraz na minimálne alebo žiadne emisie. Nasadzovanie autobusov s elektrickým pohonom je podporované štátmi a samosprávami, preto v tejto oblasti možno očakávať širšie nasadenie vozidiel s elektrickým pohonom.

Vývoj batérových technológií sa ubera viacerými smermi. Tým hlavným je zmena technológie samotnej batérie s cieľom zvýšiť jej energetickú hustotu, schopnosť rýchleho nabíjania bez poškodzovania batérie a zvýšenia jej bezpečnosti. Okrem toho sa hľadajú možnosti, ako predĺžiť životnosť batérie inými prostriedkami. Životnosti batérie neprospevia nabíjanie a vybíjanie vysokými prúdmi. EV však počas svojej prevádzky vyžaduje zvýšený odber energie pri akcelerácii a tiež schopnosť rýchlo prijať energiu pri brzdení. Takéto jazdné cykly sú typické najmä pre mestskú prevádzku. Pri dlhodobo opakovanom zaťažovaní batérie takýmto spôsobom môže dôjsť k zhoršeniu jej vlastností už po veľmi krátkom čase.

Aby sa batéria uchránila od vysokých prúdov, ale zároveň sa zachovala možnosť vozidla rýchlo akcelerovať, možno do zásobníka energie vo vozidle pridať zariadenie, ktoré je schopné prijať v krátkom čase veľké množstvo energie a následne ju aj poskytnúť. Jedným z riešení je doplnenie trakčnej batérie o superkapacitor (SC). SC sa vyznačuje vysokým merným výkonom, čo znamená, že dokáže prijať aj dodať vysoký prúd. Na druhej strane má nízku mernú energiu, takže na predĺženie dojazdu vozidla sa veľmi neprejaví. Zabezpečenie dostatočného dojazdu vozidla zostáva úlohou batérie. Z toho vyplýva, že je výhodné, aby vysoké špičky výkonu



pokrýval SC a dlhodobé zaťaženie batéria.

SC má výrazne nižší vnútorný odpor než priemerná batéria, čo umožní znížiť celkové vnútorné straty zásobníka energie a zvýšiť jeho účinnosť. Použitie kombinácie batéria-SC má však aj svoje nevýhody. Napätie SC sa výrazne mení na základe stavu nabitia SC, čo vyžaduje použitie prídavného DC/DC meniča, cez ktorý by sa SC pripájal na DC zbernicu. Okrem toho zaradenie SC a ďalšieho meniča do trakčného reťazca vozidla zvyšuje náklady a hmotnosť vozidla. Veľkosť batérií a SC je preto nutné optimalizovať pre danú aplikáciu.

Článok sa zaoberá návrhom zásobníka energie pre midibus určený pre MHD. Zásobník je zostavený z batérií a superkapacitora, a je doplnený o algoritmus prerozdelenia výkonov medzi danými komponentmi. Model batérie obsahuje aj model jej starnutia. Overenie vlastností zásobníka bolo robené simulačne v prostredí Matlab. Súčasťou návrhu bola aj metodika na optimalizáciu veľkostí batérie a superkapacitora.

## II. SÚČASNÝ STAV PROBLEMATIKY

Existuje viacero možných topológií zapojenia batérie a SC [1]. Zlepšenie účinnosti a výkonu sa dosahuje už pri paralelnom zapojení batérie a SC bez použitia DC/DC meniča, potenciál SC však zostáva vo veľkej miere nevyužitý. Pridaním DC/DC meniča sa dosiahne možnosť prevádzky SC v širokom rozsahu napätí a dá sa tak využívať väčšie percento kapacity SC [2]. Použitie obojsmerného DC/DC meniča pre SC sa javí ako dobrý kompromis medzi dosiahnutím želaného efektu a nákladmi na systém. Vzhľadom na tieto výhody je spomínané riešenie momentálne najčastejšie využívanou topológiou [3]. Jednou z navrhnutých možností je aj použitie dvojice komplementárnych tranzistorov namiesto DC/DC meniča [4]. V prevádzke by bola buď batéria alebo SC. Prinieslo by to zlepšenie účinnosti zásobníka, napätie DC medziobvodu by však prudko kolísalo vzhľadom na napätie SC, čo by vyžadovalo zložitejší algoritmus riadenia motorového meniča.

Toky energie medzi batériou, SC a motorom sú riadené manažmentom energie. Bolo navrhnuté veľké množstvo rôznych prístupov, ktoré boli založené na optimalizačných technikách, fuzzy logike, neurónových sieťach a pod. Jedným zo spôsobov je použitie minimalizačnej funkcie [5]. Tento návrh pokrýva široký rozsah výkonov EV, ale SC je niekedy vybíjaný až na nulu a pomerne výrazne zaťažovaný. Algoritmus manažmentu energie je dobre popísaný v [6] a vychádza z riadenia stavu nabitia SC vzhľadom na kinetickú energiu vozidla. Dobrou možnosťou je použitie fuzzy logiky [7]. Pri tomto návrhu však prevažovalo vybíjanie nad nabíjaním, preto dochádzalo k stavom, kedy bol SC výrazne vybitý aj pri nízkych rýchlostiach. Ďalšou možnosťou je, že by SC plnil funkciu hornopriepustného filtra, kedy by záťaž s nízkou frekvenciou kryla batéria a záťaž s vysokou frekvenciou SC [8].

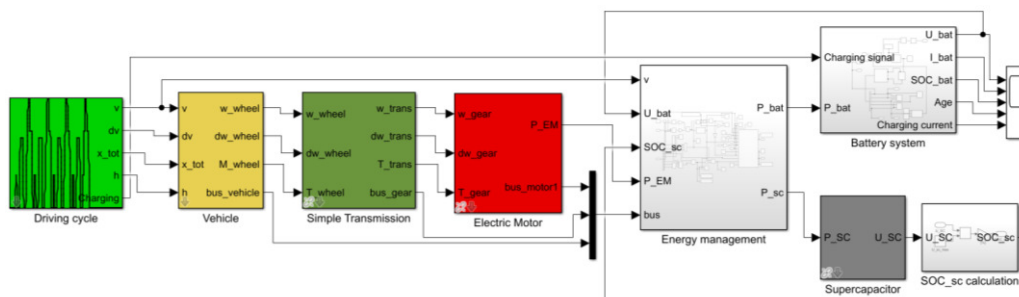
Aby bolo možné overiť správnosť návrhu simuláciou, je potrebné mať čo najpresnejší model batérie. Existuje viacero rôznych modelov batérií. V práci sa využívali modely batérií z programu MATLAB/Simulink. Jedná sa o blok „Equivalent Circuit Battery“ zostavený podľa [9] alebo „battery“ z knižnice Simscape [10], [11]. Pri „Equivalent Circuit Battery“ je napätie batérie určené pomocou tabuľkových hodnôt, pričom tabuľky sú funkciami stavu nabitia a teploty batérie. Výhodou tohto modelu je, že si používateľ môže nastaviť parametre batérie prostredníctvom vybíjacej charakteristiky batérie. Model bloku „battery“ z knižnice Simscape je o niečo komplikovanejší, vybíjacia charakteristika je popísaná rovnicami. Je tu zahrnutý aj tepelný model [12], [13] a model starnutia batérie [14]. Práve pre možnosť simulácie starnutia batérie sa v práci využíval tento blok. Okrem toho bol v simulácii použitý model superkapacitora podľa [15]. Zároveň bola využitá topológia s obojsmerným DC/DC meničom a manažmentom energie na základe kinetickej energie vozidla, ktorý umožní presné rozdelenie výkonu medzi oba komponenty.

## III. MATEMATICKÝ MODEL MIDIBUSU

Model vozidla so zásobníkom energie bol vytvorený v QSS Toolbox-e [16]. QSS TB poskytuje možnosť jednoduchého zostavenia modelu rôznych pohonných systémov vozidla a odhad jeho spotreby pri rôznych jazdných cykloch, pričom nie je náročný na výpočtový výkon počítača. QSS TB využíva kvázistatickú metódu, pri ktorej sa uvažuje len s mechanickými prechodnými dejmi. Výpočet rieši predovšetkým energetickú bilanciu jednotlivých komponentov trakčného reťazca, pri ktorej sú komponenty reprezentované statickými charakteristikami. Využíva sa tzv. spätný výpočet, t.j. vychádza sa z jazdného cyklu. Ten definuje priebeh rýchlosti vozidla v čase. Pri simulácii sa celý jazdný cyklus rozdelí na krátke časové úseky – výpočtové kroky, ktorých dĺžka sa volí tak, aby sa hodnota trakčných odporov medzi dvoma susednými krokmi výrazne nemenila. Na základe predpísaného priebehu rýchlosti sa počíta zrýchlenie vozidla, prejdená dráha a príp. zaradený prevodový stupeň. Na ich

základe a na základe parametrov jednotlivých komponentov trakčného reťazca sa počíta potrebný výkon motora a energia na jazdu.

Model trakčného reťazca pozostáva z niekoľkých subsystémov (obr. 1). Blok „Driving cycle“ umožňuje výber niektorého z bežne používaných jazdných cyklov, ale ponúka aj možnosť načítať vlastné dáta o jazdom cykle namerané na reálnej trase. Blok je upravený tak, že umožňuje načítať aj údaje o nadmorskej výške a čase nabíjania. V ďalšom bloku „Vehicle“ sa počítajú trakčné odpory podľa zadaných parametrov vozidla. Blok „Single Transmission“ predstavuje model jednostupňovej prevodovky. V bloku „Electric Motor“ sa počíta požadovaný elektrický príkon motora z uhlovej rýchlosti a momentu motora. Účinnosť motora sa tu určuje na základe mapy účinnosti. Na výstupe je požadovaný príkon, ktorý musí poskytnúť zásobník energie. Tento príkon sa manažmentom energie vhodne rozdelí medzi batériu a SC.

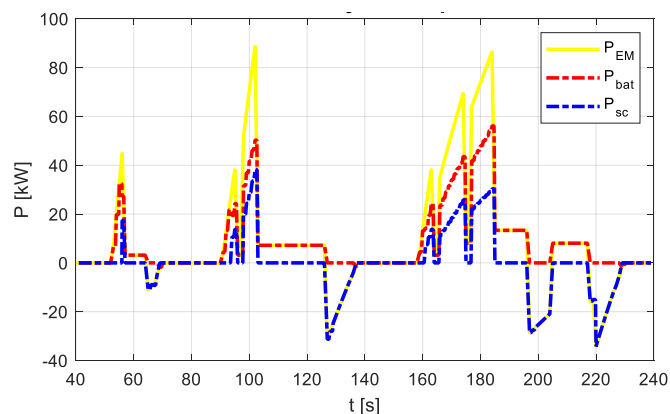


Obr. 1 Model midibusu zostavený pomocou QSS Toolbox-u a model zásobníka s manažmentom energie.

#### A. Manažment energie

Manažment energie funguje na princípe riadenia stavu nabitia SC vzhľadom na kinetickú energiu vozidla. To umožňuje vysoké využitie energie získanej pri rekuperačnom brzdení a kedykoľvek poskytuje energiu aj pre veľké zrýchlenia. Princíp spočíva vo vybíjaní SC takým spôsobom, aby v ňom zostalo dosť miesta na absorbovanie energie z rekuperačného brzdienia. SC sa vybíja len takým výkonom, aby pri zabrzdení na nulovú rýchlosť bol SC plne nabitý. Touto požiadavkou sa zabráni nadmernému vybitiu, a preto je možné zahájiť nový jazdný cyklus aj pri nárokoch na vysoké zrýchlenia.

Blok „Energy management“ obsahuje pomerne komplikované výpočty a algoritmus, ktorý je podrobne popísaný v [6]. Výsledkom je, že pri dlhodobej záťaži sa energia odoberá z batérie a pri krátkodobej špičkovej záťaži už spolupracuje aj SC, pričom pri rekuperácii pracuje len SC. Príklad rozdelenia výkonov pri jazdom cykle ECE je na obr. 2.



Obr. 2 Rozdelenie výkonov medzi batériu a SC pri jazdom cykle ECE.

#### B. Zásobník energie

Model zásobníka energie sa skladá z dvoch hlavných častí, modelu batérie zo Simscape a superkapacitára z QSS toolbox-u. Ku nim boli doplnené ďalšie bloky, ktoré slúžia na riadenie nabíjania a vybíjania batérie, zisťovanie stavu nabitia SC, výpočet dojazdu vozidla a výpočet vnútorných strát. Viac podrobností je uvedených v [17].

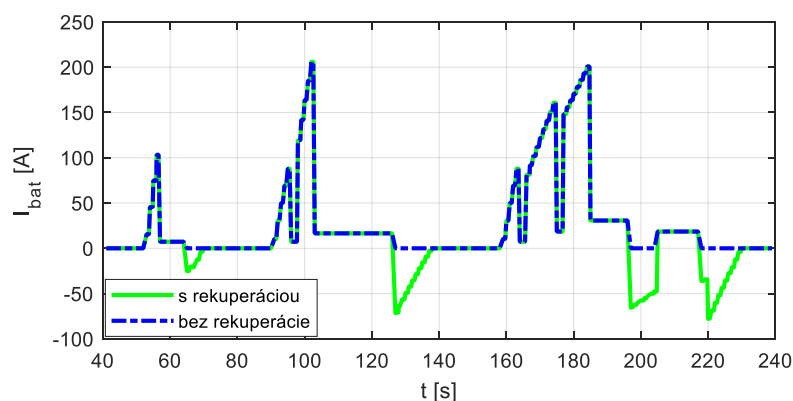
### IV. SIMULÁCIA MODELU MIDIBUSU

Cieľom simulácií bolo zistiť, do akej miery je použitie SC výhodné. Sledovala sa zmena

starnutia batérie, spotreba energie, vnútorné straty, dojazd vozidla a ďalšie faktory. Pri simulácii sa používali dva jazdné cykly, ECE a NEDC, pričom maximálna rýchlosť pri NEDC bola obmedzená na 90 km/h. Najprv bola urobená simulácia, kde bola zásobníkom energie len batéria. Simulovala sa jazda s rekuperáciou a bez rekuperácie. Výsledkom bolo množstvo energie, ktorá sa dá ušetriť rekuperačným brzdením. V ďalších simuláciách bola zásobníkom kombinácia batérie a SC. Celková kapacita SC bola zložená z niekoľkých paralelne zapojených superkapacitorových modulov, ktorých počet  $N_{sc}$  sa dal meniť vzhľadom na jazdný cyklus. V simuláciách sa sledoval aj vplyv maximálnej hodnoty prúdu odoberaného z batérie.

#### A. Simulácie s rekuperáciou a bez rekuperácie

Najprv bol simulovaný cyklus ECE a NEDC pri napájaní len z batérie, SC sa nepoužil. Na obr. 3 sú priebehy prúdov pre cyklus ECE s rekuperáciou a bez nej. Výsledky spotreby energie ukázali, že bez rekuperácie bola spotreba 24,85 kWh a s rekuperáciou 18 kWh. Rekuperáciou sa pri tomto jazdnom cykle tak dá dosiahnuť úspora až 27,5 %. Pri jazdnom cykle NEDC spotreba pri prevádzke bez rekuperácie dosiahla 59,56 kWh a bez rekuperácie 47,3 kWh. Úspora v tomto prípade činila 20,57% energie.

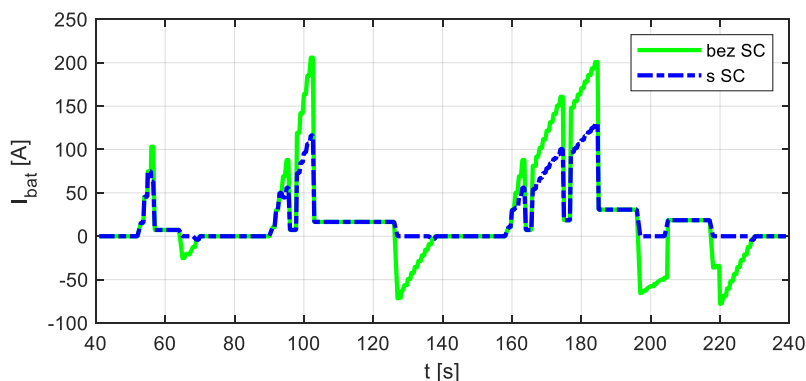


Obr. 3 Priebeh prúdu pri napájaní len z batérie pre jazdný cyklus ECE.

#### B. Vplyv veľkosti maximálneho prúdu batérie

Následne boli vykonané simulácie s využitím SC. Menil sa maximálny prúd batérie  $I_{bat,max}$ , ktorý nadobúdal hodnoty 60 A, 80 A, 100 A, 120 A a 200 A (0,3C, 0,4C, 0,5C, 0,6C a 1C). Z predchádzajúcich simulácií vyplynulo, že celková kapacita SC by mala byť aspoň 24,61 F. Keďže kapacita jedného SC modulu je 6 F, použilo sa 5 modulov s celkovou kapacitou 30 F. Sledovalo sa starnutie batérie  $age$  (v počte ekvivalentných plných cyklov), percento využitia kapacity SC  $C_{used}$ , vnútorné straty  $E_i$  a dojazd AER. Výsledky simulácií sú zhrnuté v Tab.1.

Prúdy pretekajúce batériou boli pri použití SC výrazne nižšie, čomu zodpovedá aj spomalenie starnutia, ktoré bolo až takmer o polovicu pomalšie. Špičky pri akcelerácii sú podstatne zmenšené a špičky pri rekuperačnom brzdení dokonca úplne odstránené (obr. 4). Zmena maximálneho prúdu batérie nepriniesla výraznejšie zmeny na životnosti, no treba poznamenať, že sa pracovalo s prúdmi do hodnoty 1C. Dojazd vozidla sa predĺžil len mierne. Zaujímavým ukazovateľom je percento využitia SC. Obmedzením  $I_{bat,max}$  na nižšie hodnoty sa začala kapacita SC využívať v širšom rozsahu a dochádzalo ku hlbšiemu vybíjaniu SC.



Obr. 4 Priebeh prúdu batérie bez a s použitím SC pre jazdný cyklus ECE.

Pri  $I_{bat\_max} = 80A$  už kapacita SC nepostačovala nárokom a stav jeho nabitia klesol až na dolné obmedzenie. Kapacita SC bola zvýšená, inak by zvyšok výkonu musela dodávať batéria a bola by počas tohto úseku preťažovaná. Z uvedených výsledkov vyplýva, že je možné znižovať odoberaný prúd z batérie, ale vyžaduje si to väčšiu kapacitu SC.

Použitím SC sa sledovalo aj zníženie vnútorných strát zásobníka energie. So znižujúcim sa  $I_{bat\_max}$  sa však vnútorné straty začali znova zvyšovať. Dôvod je práca SC pri nižších hodnotách stavu nabitia. Na dosiahnutie rovnakého výkonu pri nižšom stave nabitia potrebujeme z SC odoberať vyšší prúd, čím narastajú aj jeho vnútorné straty.

Tabuľka 1 Výsledky simulácií pre cyklus ECE

ECE	$N_{sc}$	$age$	$C_{used}$ [%]	$E_i$ [MJ]	AER [km]
Bez rekuperácie	0	0,11	0	9,813	112,8
S rekuperáciou	0	0,08327	0	8,544	155
$I_{bat\_max} = 200 A$	5	0,04292	63,91	7,805	158,3
$I_{bat\_max} = 120 A$	5	0,04288	64,79	7,817	158,4
$I_{bat\_max} = 100 A$	5	0,04292	82,59	9,133	160,1
$I_{bat\_max} = 80 A$	6	0,04349	77,94	9,466	160,1
$I_{bat\_max} = 60 A$	7	0,04263	90,11	13,24	163,2

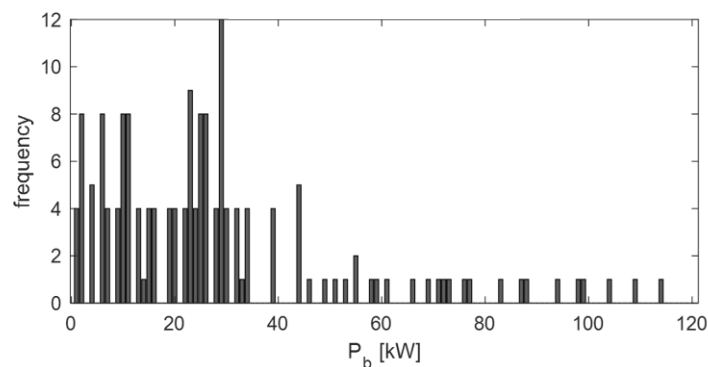
Jedným zo spôsobov ako riešiť manažment energie by mohla byť rekuperácia malých prúdov do batérie a veľkých do SC. Výhodou by bolo, že by stačila menšia celková kapacita SC. Simulácie ukázali, že režim prevádzky by sice mal nepriaznivý vplyv na životnosť batérie, starnutie by však bolo stále pomalšie, ako keby sa SC nepoužil vôbec.

### C. Možnosti vylepšenia modelu a simulácií

Na základe uvedených simulácií možno konštatovať, že použitie SC v zdrojovom reťazci vozidla dokáže výrazne znížiť namáhanie batérie. Presný údaj o starnutí bol získaný len na základe modelu starnutia bloku „Battery“ v Simscape, ktorý má ešte rezervy. Ideálnou možnosťou ako zistiť mieru starnutia batérie, je jej fyzické vybíjanie na testeri batériových článkov, kde by sa zisťovala rýchlosť starnutia pri rôznych priebehoch prúdu. Výsledky by sa neskôr použili na úpravu modelu batérie a optimalizáciu zloženia zostavy batéria-SC.

### D. Optimalizácia veľkosti komponentov zásobníka energie

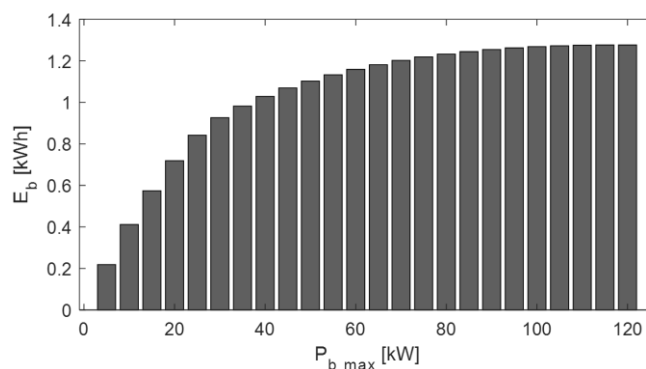
Každý jazdný cyklus má iný priebeh rýchlosti, iné veľkosti akcelerácie a brzdenia, z toho vyplývajúce rôzne požiadavky na veľkosť výkonu. Nároky na požadovaný výkon sú hlavným faktorom pri voľbe veľkosti komponentov zásobníka. Pri ich optimalizácii pomôže analýza dosahovaných výkonov počas daného jazdného cyklu. Analýza sa robí na základe výkonového histogramu, ktorý ukáže, ako často sa dané hodnoty výkonu počas cyklu vyskytovali. Výsledky ukázali, že maximálny brzdny výkon pri ECE nepresahuje 40 kW, pričom pri NEDC sa blíži až k 115 kW (obr. 5). Z histogramov možno konštatovať, že pri oboch cykloch sú najfrekvencovanejšie brzdne výkony približne do 35 kW.



Obr. 5 Histogram brzdneho výkonu pre jazdný cyklus NEDC.

Ak by sa veľkosť komponentov zásobníka zredukovala, dali by sa rekuperovať len menšie výkony. Preto je vhodné vedieť, aké množstvo energie dokáže zásobník so zníženým

maximálnym výkonom zrekuperovať počas jedného cyklu (obr. 6). Zvyšok energie by sa maril v mechanických brzdách.



Obr. 6 Veľkosť zrekuperovanej energie vzhľadom na maximálny brzdný výkon pri cykle NEDC.

## V. ZÁVER

Článok prezentuje možnosť využitia kombinovaného zásobníka elektrickej energie pre elektrický midibus. Zásobník obsahuje batériu a superkapacitor a je doplnený o systém manažmentu energie, ktorý riadi rozdelenie energie medzi obe časti zásobníka. Analýza bola robená simulačne v prostredí Matlab/Simulink, pričom model batérie zohľadňoval aj jej starnutie. Simulácie potvrdili, že pri použití superkapacitora sa dajú prúdy pretekajúce batériou výrazne znížiť. Týka sa to hlavne špičkových prúdov batérie pri akcelerácii a brzdení vozidla. Hoci predĺženie dojazdu vozidla sa zaradením superkapacitora výrazne nezvyšilo, dosiahlo sa spomalenie starnutia batérie, ktoré bolo podľa výsledkov približne o polovicu pomalšie. Vďaka tomu batéria vydrží v prevádzke dlhšiu dobu bez nutnosti jej výmeny.

Manažment energie zásobníka vychádzal z využitia kinetickej energie vozidla a zabezpečil, aby bol superkapacitor po dobrzdení vozidla plne nabitý a pripravený na následný rozjazd vozidla. Manažment bol doplnený o algoritmus, ktorý okrem kinetickej energie dokázal zohľadniť aj potenciálnu energiu vozidla pri jazde po členitom teréne.

V rámci analýzy sa overovalo správanie zostavy zásobníka pre rôzne rozdelenie prúdov medzi batériu a superkapacitor. Sledovali sa predovšetkým vnútorné straty zásobníka. Čím bol prúd batérie nižší, tým väčšia kapacita SC bola využívaná, a tým väčší prúd pretekal cez SC. Napriek tomu, že SC má nižší vnútorný odpor než batéria, pri veľkých prúdoch cez SC rastú výrazne vnútorné straty zostavy. Z toho vyplýva, že je možné znižovať prúd odoberaný z batérie, ale vyžaduje si to väčšiu kapacitu SC.

Testovala sa aj možnosť čiastočného rekuperovania energie batériou. V tomto prípade postačuje použitie menšej kapacity SC, ale efekt predĺženia životnosti batérie už nie je taký výrazný.

V práci bola načrtnutá metodika optimalizácie veľkosti komponentov zásobníka vzhľadom na konkrétny jazdný cyklus pomocou histogramov brzdných výkonov.

Najvýraznejším prínosom kombinovaného zásobníka energie je predĺženie životnosti batérie. Tým, že si batéria udrží dostatočnú kapacitu dlhšiu dobu, nie je nutná jej výmena a recyklácia. Nakoľko energetické nároky na recykláciu batérií pri súčasných technológiách sú podstatne vyššie ako pri výrobe novej batérie, možnosť prevádzkovania pôvodnej batérie vo vozidle je výrazným prínosom k ochrane životného prostredia.

## POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-15-0750.

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-18-0436.

## LITERATÚRA

- [1] A. Khaligh, L. Zhihao, "Battery, ultracapacitor, fuel cell, and hybrid energy storage systems for electric, hybrid electric, fuel cell, and plug-in hybrid electric vehicles: state of the art". *IEEE Transactions on Vehicular Technology*, 2010; 59(6): pp. 2806–2814.
- [2] L. Gao, R. A. Dougal, and S. Liu, "Power enhancement of an actively controlled battery/ultracapacitor hybrid," *IEEE Transaction on Power Electronics.*, vol. 20, no. 1, pp. 236–243, Jan. 2005.
- [3] L. Kouchachvili, W. Ya'ici, E. Entchev, "Hybrid battery/supercapacitor energy storage system for the electric vehicles", *J. Power Sources* 374 (2018) pp. 237–248.

- [4] L. Shuai, K. A. Corzine, M. A. Ferdowsi, "A new battery/ultracapacitor energy storage system design and its motor drive integration for hybrid electric vehicles", *IEEE Transactions on Vehicular Technology*, 2007; 56(4): pp. 1516–1523.
- [5] M.-E. Choi, S.-W. Seo, "Robust energy management of a battery/supercapacitor hybrid energy storage system in an electric vehicle", in: *Electric Vehicle Conference (IEVC)*, 2012 IEEE International, March 2012, pp. 1-5.
- [6] J. Armenta, C. Núñez, N. Visairo, I. Lázaro, "An advanced energy management system for controlling the ultracapacitor discharge and improving the electric vehicle range", *J. Power Sources* 284 (2015) pp. 452–458.
- [7] L. Rosario, P.-K. Luk, "Applying management methodology to electric vehicles with multiple energy storage systems", in: *International Conference on Machine Learning and Cybernetics*, 2007, vol. 7, Aug 2007, pp. 4223 - 4230.
- [8] E. Schaltz, A. Khaligh, P. O. Rasmussen, "Influence of battery/ultracapacitor energy-storage sizing on battery lifetime in a fuel cell hybrid electric vehicle". *IEEE Transactions on Vehicular Technology*, 2009; 58(8), pp.3882–3891.
- [9] R. Jackey, "A Simple, Effective Lead-Acid Battery Modeling Process for Electrical System Component Selection." *SAE Technical Paper* 2007-01-0778. doi:10.4271/2007-01-0778, 2007.
- [10] O. Tremblay, L. A. Dessaint, and A. I. Dekkiche, "A generic battery model for the dynamic simulation of hybrid electric vehicles," in *Proc. IEEE Veh. Power Propulsion Conf. (VPPC 2007)*, pp. 284–289.
- [11] O. Tremblay, L. A. Dessaint, "Experimental Validation of a Battery Dynamic Model for EV Applications." *World Electric Vehicle Journal*. Vol. 3, May 13–16, 2009.
- [12] L. H. Saw, K. Somasundaram, Y. Ye, A. A .O. Tay, "Electro-thermal analysis of Lithium Iron Phosphate battery for electric vehicles." *Journal of Power Sources*. Vol. 249, pp. 231–238.
- [13] C. Zhu, X. Li, L. Song, and L. Xiang, "Development of a theoretically based thermal model for lithium ion battery pack." *Journal of Power Sources*. Vol. 223, pp. 155–164.
- [14] N. Omar, M. A. Monem, Y. Firouz, J. Salminen, J. Smekens, O. Hegazy, H. Ghaulous, G. Mulder, P. Van den Bossche, T. Coosemans, and J. Van Mierlo, "Lithium iron phosphate based battery — Assessment of the aging parameters and development of cycle life model." *Applied Energy*, Vol. 113, January 2014, pp. 1575–1585.
- [15] L. Guzzella, C. H. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*, Springer Verlag, Berlin, 2004.
- [16] QSS toolbox. ETH Zürich – Swiss Federal Institute of Technology. <https://idsc.ethz.ch/downloads>.
- [17] D. Bodnár, *Predikcia spotreby elektrického midibusu*. Diplomová práca. FEI TU Košice, 2021.

# Prototypovanie riadenia dronov v zdieľanej webovej virtuálnej realite

<sup>1</sup>Dávid GULA, <sup>2</sup>Štefan KOREČKO, <sup>3</sup>Branislav SOBOTA

Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská Republika

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

**Abstrakt** – Súčasťou vývoja zariadení je aj tvorba a testovanie ich prototypov, ktorá je častokrát zdĺhavá a finančne náročná. Je potrebné vytvoriť a postaviť fyzický prototyp, ktorý sa potom testuje v reálnom prostredí a vyhodnocuje pomocou rôznych metrík a senzorov. Taktiež je nutné zaručiť bezpečnosť účastníkov, keďže pri poruche reálneho zariadenia môže dôjsť k materiálnym, či dokonca zdravotným škodám. V tomto príspevku predstavujeme zdieľané virtuálne prostredie, v ktorom je možné overovať funkčnosť prototypov zariadení využitím technológií virtuálnej reality. Prostredie je možné spustiť jednoducho vo webovom prehliadači a podporuje pripojenie viacerých používateľov naraz prostredníctvom počítača, smartfónu, tabletu, či prilby pre rozšírenú a virtuálnu realitu. Účelom prostredia je prototypovanie riadenia dronov v ňom s ohľadom na bezpečnosť používateľov. Prezentovaný je aj scenár, kde dron roznáša materiál v prostredí staveniska tak, aby nedošlo k ohrozeniu používateľov.

**Kľúčové slová** – ovládanie dronov, prototypovanie, virtuálna realita, zdieľané prostredie

## I. ÚVOD

Virtuálna realita je oblasť, ktorá existuje už dlhšiu dobu, no až v poslednom desaťročí sa výrazne rozširuje. Najčastejšie nájde využitie v medicíne (rehabilitácie, liečenie fobií, svalových a nervových porúch, atď.), vzdelávaní a trénovaní na diaľku (vzdelávanie žiakov a študentov, trénovanie zamestnancov, pracovníkov a lekárov, simulácia krízových situácií, atď.).

Okrem toho je možné virtuálnu realitu využiť aj pri návrhu a testovaní prototypov rôznych zariadení. Táto fáza vývoja produktu je veľmi dôležitá, avšak zdĺhavá a finančne náročná. Taktiež si vyžaduje vytvoriť reálny prototyp produktu a overovať ho fyzicky, za pomoci prístrojov a prítomnosti odborníkov. Virtuálna realita tu ponúka možnosť vykonať testovanie vo virtuálnom prostredí, ktoré je možné jednoducho nakonfigurovať podľa potreby, vytvoriť rôzne scenáre a overovať tak funkčnosť prototypu v rôznych podmienkach. To všetko je možné realizovať bez nutnosti tvorby reálneho zariadenia. Virtuálne prostredia ale nedokážu simulovať všetky aspekty reálneho sveta, preto nie je možné reálne testovanie úplne nahradiť.

Príkladom zariadení, pre ktoré je veľmi dôležité ich otestovanie pred nasadením do prevádzky, sú lietajúce drony, najmä autonómne ovládané. V dnešnej dobe sú veľmi rozšírené komerčné drony, keďže sú lacné, dostupné a nájdu využitie v mnohých oblastiach, od fotografovania a prieskumu, cez doručovanie balíčkov, až po asistenciu počas rôznych nešťastí a hľadania ranených [1]. Veľkou výzvou v tejto oblasti je vytvoriť spoľahlivý riadiaci softvér, ktorý dokáže autonómne ovládať dron a splniť rôzne požiadavky počas plnenia úloh. Jedná sa napríklad o bezpečnosť prevádzky, rýchlosť a efektivitu plnenia úloh a schopnosť reagovať na okolité prostredie.

Ako veľmi sľubná platforma sa v tomto kontexte javí zdieľaná webová virtuálna realita, kde virtuálne prostredia (scény) sú zobrazené vo webovom prehliadači. A jedno takéto prostredie môže byť naraz zdieľané viacerými používateľmi. Je pri tom možná aj vzájomná interakcia používateľov na rôznej úrovni.

V tomto príspevku predstavujeme pilotnú implementáciu virtuálneho prostredia, použiteľného pre testovanie a overovanie prototypov riadiacich softvérov lietajúcich dronov. Virtuálne prostredie pozostáva zo scény staveniska, v ktorej sa nachádza dron aj používateľa. Pre prostredie bol implementovaný aj scenár, teda činnosť, a to pre dron aj používateľov. Dron opakovane navštevuje vopred definované miesta, kde vždy pristane a na chvíľu zotrúva. Úlohou používateľov je nájsť v prostredí určité pozície.

Prostredie aj scenár boli vytvorili v systéme *LIRKIS G-CVE* [3], [4] pre zdieľanú webovú virtuálnu realitu, postavenom na báze rámca *A-Frame*<sup>1</sup>. Pre pohyb dronu je použitá fyzikálna

<sup>1</sup><https://aframe.io/>

simulácia v reálnom čase. Pre jeho autonómne riadenie bol implementovaný riadiaci systém, ktorý dron ovláda tak, aby bol schopný presúvať sa medzi pozíciami, prípadne na danej pozícii bezpečne pristáť a bezpečne z nej znova vzlietnuť. Tieto manévry vykonáva tak, aby nedošlo k ohrozeniu používateľov v jeho blízkosti. Na získavanie informácií o pripojených používateľoch sme použili komponent pre zdieľanie pozícií pripojených entít v prostredí [2], označovaný ako Enhanced Client Access (ECA) [3], [4].

## II. SÚVISIACE PRÁCE

Existuje viacero rôznych systémov, ktoré podporujú virtuálnu realitu a zdieľané virtuálne prostredie pre viacerých používateľov. Väčšina z nich sa však nezameriava na overovanie prototypov zariadení. Okrem toho sú taktiež prevažne vytvárané pomocou herných rámcov, najčastejšie *Unity*, teda nejedná sa o webové virtuálne prostredia.

Systém, ktorý má najbližšie k nášmu zdieľanému prostrediu je kolaboratívne VR prostredie pre návrh produktov [5]. Systém umožňuje dizajnérom a inžinierom pracovať s prototypmi produktov a zostavovať ich v zdieľanom virtuálnom prostredí, avšak nesmú byť príliš zložitý. Vytvorený je pomocou rámca *Unity* a rozšírenia *Virtual Reality Toolkit*. Podporuje iba virtuálne prilby *HTC Vive*. Pomocou ovládačov používateľa manipulujú s objektami. Nami implementovaný systém je narozdiel od tohto realizovaný vo webovom prostredí a podporuje možnosť pripojiť sa aj imerzívne, prostredníctvom prilby pre virtuálnu realitu, ale aj pomocou bežných zariadení. Taktiež je ho možné jednoducho spustiť v prehliadači, bez nutnosti inštalácie softvéru. Ďalšou výhodou nášho systému je možnosť testovať aj funkcionality prototypov, narozdiel od prezentovania iba vzhľadu a 3D objektu.

Ďalšie spomínané systémy sa zameriavajú na simuláciu lietajúceho dronu vo virtuálnom prostredí a jeho interakciu s objektmi.

Jedným z takýchto systémov je simulátor *DroneVR* [6]. Implementovaný je vo webovom prostredí pomocou vizualizačného rámca *ThreeJS*. Jeho cieľom je minimalizovanie rizika nárazu dronu ovládaného buď manuálne, alebo pomocou autopilota, do budovy alebo iného objektu. Princípom je modelovanie reálneho sveta a budov vo virtuálnom prostredí, v ktorom sa človek môže naučiť ovládať zariadenie v konkrétnom simulovanom prostredí, ktoré zodpovedá reálnemu prostrediu. Dáta o budovách a ich parametroch systém čerpá zo systému *OpenStreetMap*, podľa ktorého vytvorí virtuálne prostredie. Do nášho systému je nutné objekty prostredia pridávať ručne, je ale možné takúto funkcionality do systému doplniť. Ovládanie dronu v systéme *DroneVR* je realizované buď manuálne, alebo pomocou integrovaného autopilota (v tomto móde sa dron vyhýba prekážkam). Avšak v prípade riadenia autopilotom sa dron pohybuje iba v rovnakej výške nad zemou a je schopný sa pohybovať a otáčať sa iba v ôsmich rôznych smeroch. Náš systém simuluje dron, ktorý sa pohybuje v akomkoľvek smere a výške. Taktiež simulujeme štart a pristávanie.

Ďalším simulátorom pre drony je systém *AirSim* [7]. Je implementovaný v hernom rámci *Unreal Engine*, vďaka čomu je veľmi realistický a graficky detailný. Umožňuje integrovať rôzne ovládacie firmvéry pre drony, ktoré sa používajú aj v reálnych strojoch. Okrem toho simuluje rôzne senzory, ako napríklad akcelerometer, gyroskop, barometer, magnetometer a GPS. Taktiež umožňuje simulovať rôzne efekty počasia, ktoré vplývajú na let dronu. Tento simulátor poskytuje realistickejšie spracovanie dronu, ako náš systém, avšak neumožňuje používateľom interagovať s dronom a taktiež nepodporuje viacerých používateľov v rovnakom čase.

Simulátor *Flightmare* [8] sa od *AirSim* líši najmä v tom, že autori oddelili vizualizačný rámec od fyzikálnej simulácie. Dôsledkom je zvýšenie výkonnosti simulácie použitím paralelného programovania. Na vizualizáciu používajú rámec *Unity*. Taktiež autori implementovali rozhranie pre nástroje strojového učenia, takže sa prostredie môže použiť aj na učenie autonómne ovládaných lietajúcich dronov pomocou umelej inteligencie. Podporuje aj zariadenia pre virtuálnu realitu, vďaka ktorým sa môže používateľ pripojiť do prostredia a sledovať drony vo VR. Ani tento simulátor neumožňuje interakciu používateľ-dron. Jeho fyzikálna simulácia je ale taktiež veľmi realistická a navyše je možné simulovať viacero dronov súčasne, napríklad testovať systémy pre kolaboráciu medzi týmito strojmi.

## III. ZDIEĽANÉ VIRTUÁLNE PROSTREDIE NA PROTOTYPOVANIE RIADENIA DRONOV

Naše virtuálne prostredie a scenár boli vytvárané na platforme *LIRKIS G-CVE*, postavenej na webovom rámci *A-Frame*. Preto využívajú entitno-komponentový architektonický softvérový vzor, vďaka čomu je možné jednoducho rozširovať funkcionality pridávaním komponentov. Pre naše prostredie sme vytvorili niekoľko komponentov, ktoré rozširujú funkcionality *LIRKIS G-CVE*.



### A. Modelovanie dronu

Prvým krokom bolo vytvoriť verný model dronu, ktorý by zodpovedal svojim správaním a stavbou reálnym strojom, ktoré sú v súčasnosti využívané. Najrozšírenejšími typmi komerčných dronov sú kvadrokoptéry. Existujú dva spôsoby konfigurácie, a to konfigurácia “×” a “+”. Zvolili sme rozloženie typu “×”, keďže takýto dron je obratnejší a záťaž motorov pri vykonávaní pohybu je rovnomernejšia.

Grafický model kvadrokoptéry sme získali bezplatne, z online databázy *3D Warehouse*. Model bol dostupný vo formáte programu *SkechUp*, preto sme ho pomocou modelovacieho programu *Blender* preložili do formátu `glTF`, ktorý je použiteľný v prostredí webovej virtuálnej reality.

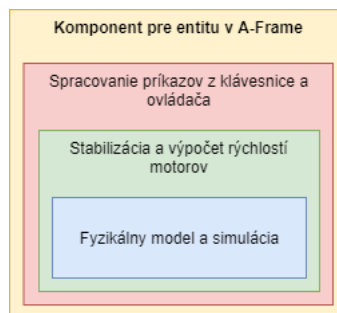
Kvadrokoptéra vykonáva svoj pohyb iba úpravou rýchlosti svojich štyroch motorov, čo z nej robí na prvý pohľad relatívne jednoduchý lietajúci stroj. Všetky motory sa po celý čas otáčajú rovnakým smerom, dva v smere a dva proti smeru hodinových ručičiek. Zmenou otáčok jednotlivých motorov vieme ovládať pohyb kvadrokoptéry v šiestich rôznych smeroch.

### B. Hierarchia riadiacich komponentov kvadrokoptéry

Pohybový systém reálnych kvadrokoptér sa skladá z niekoľkých častí:

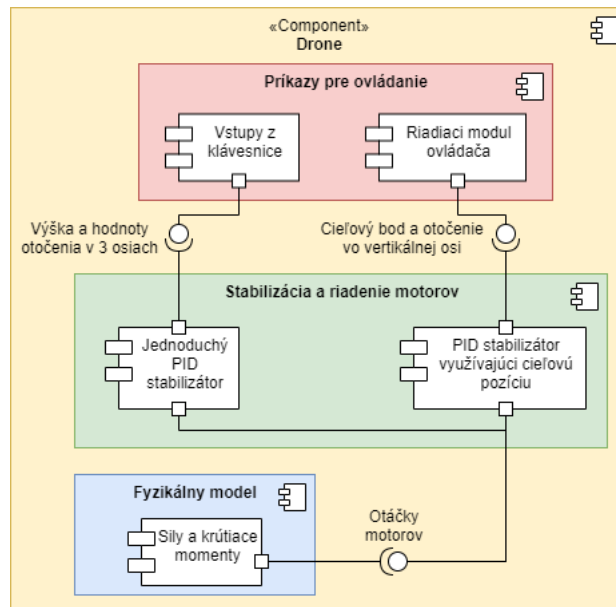
- Motory a vrtule.
- Elektronický regulátor otáčok - toto zariadenie slúži na prevod riadiaceho signálu na otáčky pre motor.
- Riadiaca doska so zabudovaným stabilizátorom - jej úlohou je reagovať na vstupy z diaľkového ovládača a podľa príkazu ovládať rýchlosti štyroch motorov.
- Autopilot - zložitejšie a drahšie kvadrokoptéry obsahujú aj autopilota, ktorý je súčasťou riadiacej dosky. Ten umožňuje kvadrokoptéram pohybovať sa autonómne.

Z hľadiska virtuálneho modelu pre nás nemá význam implementovať elektronický regulátor otáčok. Na základe spomínaných častí sme vytvorili niekoľko oddelených modulov, ktoré spolu tvoria riadiaci systém pre modelovanú kvadrokoptéru. Hierarchické usporiadanie jednotlivých častí je na Obr. 1.



Obr. 1 Hierarchia riadiacich komponentov pre model kvadrokoptéry.

Jednotlivé časti riadiaceho systému sme navrhli tak, aby boli od seba nezávislé a tým pádom ich bolo možné zamieňať a testovať tak prototypy rôznych ovládacích softvérov bez nutnosti zasahovať do nízko-úrovňového riadenia a stabilizácie dronu. Tieto časti sme vytvorili ako samostatné moduly spolu s rozhraniami, prostredníctvom ktorých medzi sebou komunikujú. Štruktúra a spôsob prepojenia modulov riadiaceho systému je zobrazený v diagrame na Obr. 2.



Obr. 2 Konceptuálny diagram riadiaceho systému kvadroptéry

Riadenie kvadroptéry pozostáva z troch hlavných častí:

- Fyzikálny model - definuje správanie sa kvadroptéry podľa reálnych fyzikálnych zákonov, teda ako má teleso reagovať na pôsobiace sily a kolízie v priestore
- PID regulátor - slúži na stabilizáciu stroja počas letu a taktiež na regulovanie otáčok motorov podľa rozdielu medzi požadovaným a súčasným stavom
- Ovládač - jeho úlohou je autonómne ovládať kvadroptéru v priestore za účelom dosiahnutia stanoveného cieľa s ohľadom na bezpečnosť používateľov

#### C. Fyzikálny rámec a fyzikálny model kvadroptéry

Na realizovanie fyzikálnych výpočtov a simulácie pôsobenia fyzikálnych síl a kolízií vo virtuálnom prostredí sme použili samostatný fyzikálny rámec. Zvolili sme si *Cannon.js* vďaka jeho podrobnej dokumentácii, jednoduchej integrácii a výkonnosti vo webovom prostredí. Prostredníctvom tohto rámca sme implementovali detekovanie kolízií medzi dronom a ostatnými statickými objektami.

Na uvedenie modelu dronu do pohybu využijeme sily, ktoré je možné v *Cannon.js* aplikovať priamo v konkrétnom bode daného objektu s určitou veľkosťou a smerom. Každý motor pôsobí dvoma silami na telo kvadroptéry: vztlakovou silou a silou točivého momentu. Ich veľkosti závisia od mnohých premenných, napríklad hustoty vzduchu, priemer a sklon lopatiek alebo vzdialenosťou motorov od ťažiska. Všetky tieto premenné sme zahrnuli do výpočtov tak, aby ich bolo možné podľa potreby nakonfigurovať. Rovnako tak aj hmotnosť a rozmery kvadroptéry je možné jednoducho nastaviť. Vďaka tomu umožňuje systém simulovať rôzne modely kvadroptér.

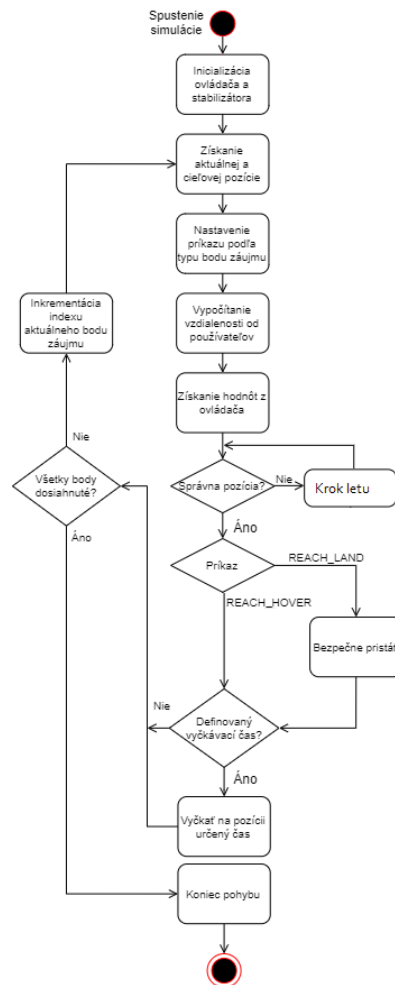
#### D. Modul stabilizácie kvadroptéry

Na tento účel sme implementovali modul využívajúci PID regulátory. Úlohou takéhoto regulátora je interpolácia medzi aktuálnou a požadovanou hodnotou. Pre stabilizáciu tak používame 6 rôznych regulátorov, každý z nich má na starosti jednu z osí, pričom pre každú os je jeden pre sledovanie pohybu a druhý pre sledovanie náklonu kvadroptéry. Modul dostane na vstupe aktuálnu a požadovanú polohu a na základe toho určí, aké otáčky je potrebné nastaviť jednotlivým motorom pre dosiahnutie tohto stavu.

#### E. Modul pre automatické riadenie

Úlohou tohto modulu je ovládať kvadroptéru autonómne tak, aby preletela všetkými bodmi záujmu a zároveň neohrozila žiadneho z používateľov vo svojom okolí. Z tohto hľadiska je zaručená bezpečnosť používateľov dvoma spôsobmi:

- 1) Kvadroptéra sa počas presúvania udržiava v bezpečnej letovej hladine, ktorú je možné ľubovoľne nakonfigurovať. Počas letu teda lieta v takej výške, kde nemôže dôjsť ku kolízii a ohrozeniu používateľov.



Obr. 3 Diagram aktivít zobrazujúci proces riadenia dronu pomocou ovládača navrhnutého pre kvadrokoptéru

- 2) Počas pristávania a štartu kvadrokoptéra sleduje minimálnu bezpečnú vzdialenosť, ktorú je možné taktiež nakonfigurovať. Ak sa niekto nachádza v tejto vzdialenosti od kvadrokoptéry počas pristávania, ovládač preruší pristávanie a vráti sa do bezpečnej letovej hladiny, kde vyčká na uvoľnenie priestoru. Ak počas štartu sa v okolí kvadrokoptéry nachádza používateľ, ovládač vyčká, kým sa priestor neuvoľní.

Činnosť ovládača a proces riadenia je znázornená v diagrame na Obr. 3. Implementovali sme ovládač tak, aby bolo možné simulovať rôzne prípady použitia dronov. Na tento účel sme implementovali príkazy, pomocou ktorých vie dron sledovať body záujmu, pristáť na konkrétnej pozícii, alebo vyčkáť vo vzduchu.

Pre tieto body sme vytvorili samostatný komponent, pomocou ktorého vieme jednoducho definovať v prostredí body záujmu pre dron. Ten si ich automaticky počas inicializácie získa a pridá do zoznamu. Vytvorili sme dva druhy bodov. *REACH* sú body, ktoré má dron iba dosiahnuť, *LAND* také, kde má aj pristáť. Každý z bodov taktiež obsahuje časový údaj definujúci, či a ako dlho má dron na danej pozícii vyčkáť.

#### F. Hlavný komponent Drone

Všetky spomínané moduly sú skombinované do jedného celku prostredníctvom hlavného komponentu *Drone*. Ten je potom možné pridať entite a urobiť tak z nej autonómne riadený dron. Hlavný komponent obstaráva inicializáciu modelu kvadrokoptéry, fyzikálneho modelu, stabilizátora a ovládača. Okrem toho poskytuje rozhranie pre konfiguráciu simulovaného modelu.

Ovládanie dronu zabezpečuje tento komponent v niekoľkých krokoch. Najprv sa získajú informácie o pozícii nasledujúceho bodu, ktorý má dron dosiahnuť a taktiež o vzdialenostiach a pozíciách všetkých pripojených používateľov vzhľadom na aktuálnu pozíciu dronu. Podľa týchto informácií určí modul pre autonómne riadenie stav a akciu, ktorú má dron vykonať. Na základe toho potom modul pre stabilizáciu vypočíta rýchlosti jednotlivých motorov. Tie sú potom použité hlavným komponentom vo fyzikálnom modeli. Tento cyklus sa opakuje neustále, kým je dron zapnutý.

### G. Virtuálne prostredie a scenár použitia

Aby sme mohli overiť použiteľnosť nami navrhnutého a implementovaného systému, vytvorili sme konkrétny scenár, ako pre používateľa, tak aj dron. Rozhodli sme sa ich zasadiť do prostredia prebiehajúcej výstavby novej budovy. Virtuálne prostredie je scénou obsahujúcou množstvo objektov, ktoré sa bežne vyskytujú na stavenisku, ako napríklad stavebný materiál, nákladné vozidlá, statické žeriavy, či kontajnery, náradie a stroje. Všetky tieto objekty<sup>2</sup> sú nastavené tak, aby cez nich používateľ, ani dron nevedeli prechádzať.

V tomto prostredí sme rozmiestnili 6 bodov tak, aby sa dron presúval medzi rôznymi pozíciami a simuloval tak konkrétny prípad použitia - presun materiálu. Do prostredia sa následne pripojili používatelia, ktorých úlohou bolo nájsť a pozbierať 6 objektov. Tie boli rozmiestnené tak, aby sa počas zbierania používateľ dron stretol a "prekážal" mu v jeho činnosti.

## IV. ZÁVER

Navrhnuté a implementované prostredie sme vytvorili na báze platformy *LIRKIS G-CVE*, vďaka čomu je možné sa do neho pripojiť využitím rôznych zariadení, nie len tých pre virtuálnu realitu. Taktiež systém podporuje pripojenie a teda aj testovanie prototypov viacerými používateľmi naraz, čo zvyšuje efektívnosť a umožňuje používateľom spolupracovať pri vývoji nového produktu. Systém bol vytvorený tak, aby boli jednotlivé moduly od seba nezávislé a bolo ich možné meniť v prípade, ak by sme chceli testovať rôzne riadiace systémy alebo fyzikálne simulácie. To všetko je implementované pomocou jedného komponentu, ktorý obstaráva všetku funkcionálnosť a je možné integrovať do rôznych prostredí. Taktiež sme docielili to, aby bolo možné jednoducho konfigurovať parametre simulácie a testovať tak riadenie konkrétnych reálnych kvadrokoptér.

Implementované prostredie sme vyskúšali aj s reálnymi používateľmi, ktorí sa v ňom pohybovali a zbierali objekty podľa scenára spomínaného vyššie. Použiteľnosť systému sme overili využitím metodológie hodnotenia použiteľnosti systému (SUS) a príslušného dotazníka. Do overenia použiteľnosti sa zapojilo 6 ľudí a celkové priemerné SUS skóre dosiahlo nadpriemernú hodnotu 81,25.

Do budúcnosti by bolo vhodné systém rozšíriť o simuláciu senzorov zariadení, pomocou ktorých reálne stroje získavajú dáta o svojom okolí. Taktiež by bolo možné implementovať detekciu prekážok vo virtuálnom prostredí a upraviť navigáciu dronu tak, aby sa im vyhýbal. Možným rozšírením a využitím systému je aj integrácia modulu pre strojové učenie, vďaka čomu by bolo možné trénovať autonómne riadenie pomocou umelej inteligencie, alebo aj kolaboráciu lietajúcich dronov za účelom dosiahnutia spoločného cieľa.

## POĎAKOVANIE

Táto publikácia bola podporená z operačného programu Integrovaná infraštruktúra v rámci projektu: Inteligentné operačné a spracovateľské systémy pre UAV, kód ITMS2014+: 313011V422 a spolufinancovaná Európskym fondom regionálneho rozvoja.

## LITERATÚRA

- [1] J. Plaza. What is the value of the european drone market? [cit. 07.01.2021]. [Online]. Available: <https://www.commercialuavnews.com/europe/value-european-drone-market>
- [2] M. Ivan, Š. Korečko, M. Hudák, and B. Sobota, "Rozšírenie možností komunikácie a správy klientov zdieľaného virtuálne-reálneho prostredia *lirkis g-cve*," in *Electrical Engineering and Informatics XI*. Faculty of Electrical Engineering and Informatics, Technical University of Košice, 2020, pp. 462–467.
- [3] M. Hudák, Š. Korečko, and B. Sobota, "Lirkis global collaborative virtual environments: Current state and utilization perspective," *Open Computer Science*, vol. 11, no. 1, pp. 99–106, 2021.
- [4] Š. Korečko, M. Hudák, B. Sobota, M. Sivý, M. Pleva, and W. Steingartner, "Experimental performance evaluation of enhanced user interaction components for web-based collaborative extended reality," *Applied Sciences*, vol. 11, no. 9, 2021. [Online]. Available: <https://www.mdpi.com/2076-3417/11/9/3811>
- [5] S. Sharma, S. Bodempudi, M. Arrolla, and A. Upadhyay, "Collaborative virtual assembly environment for product design," in *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2019, pp. 606–611.
- [6] V. T. Nguyen, K. Jung, and T. Dang, "Dronevr: A web virtual reality simulator for drone operator," in *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 2019, pp. 257–257.
- [7] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Robotics*, 2017. [Online]. Available: <https://arxiv.org/abs/1705.05065>
- [8] Y. Song, S. Naji, E. Kaufmann, A. Loquercio, and D. Scaramuzza, "Flightmare: A flexible quadrotor simulator," 2020.

<sup>2</sup>Použitie 3D modely pochádzajú z <https://majadroid.itch.io/3d-house-construction-site>.

# Návrh systému na reguláciu klímy v skleníku s využitím siete ANFIS

Ján VAŠČÁK

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

jan.vascak@tuke.sk

**Abstrakt** – Článok sa zaoberá problematikou využitia neurónovej fuzzy siete typu ANFIS na tvorbu modelu skleníka s cieľom riadenia teploty a vlhkosti jeho priestoru prostredníctvom ventilácie a zvlhčovania vzduchu. Na realizáciu takéhoto systému regulácie bol využitý jednodoskový regulátor Arduino Mega s príslušnými snímačmi. Na overenie správnosti vytvoreného modelu skleníka a spolu s ním navrhnutého riadenia bol vytvorený reálny model, na ktorom sa vykonávali experimenty.

**Keľúčovú slová** – ANFIS, Arduino, regulácia, skleníky

## I. ÚVOD

V dnešnej dobe častých klimatických zmien a náhlych teplotných výkyvov je problematické udržovať pravidelnú sezónnu úrodu. Najmä v dôsledku výskytu plesní, ktoré sa následne ťažko odstraňujú a znehodnocujú úrodu. V tomto článku popíšeme riešenie tohto problému v podobe systému na reguláciu veličín, ktoré majú priamy vplyv na kvalitu úrody, t. j. teplotu a vlhkosť [1], [2], [3].

V tomto článku sa zameriavame na vytvorenie modelu skleníka s možnosťou automatickej regulácie teploty a vlhkosti za pomoci fuzzy inferenčného systému založeného na adaptívnych sieťach, známeho pod skratkou ANFIS. Hardvérovú časť tohto riešenia pozostáva z jednodoskového počítača Raspberry Pi, mikroregulátora Arduino Mega, snímačov teploty (DHT 22, DHT 11) a snímača vlhkosti (BME 280). Tieto senzory sme využili na tvorbu dátového súboru pre model ANFIS. Taktiež sme brali do úvahy požiadavku prehľadnosti a použiteľnosti pre bežného používateľa, preto sme vytvorili aj webové rozhranie, ktoré poskytuje prehľad nad činnosťou celého systému.

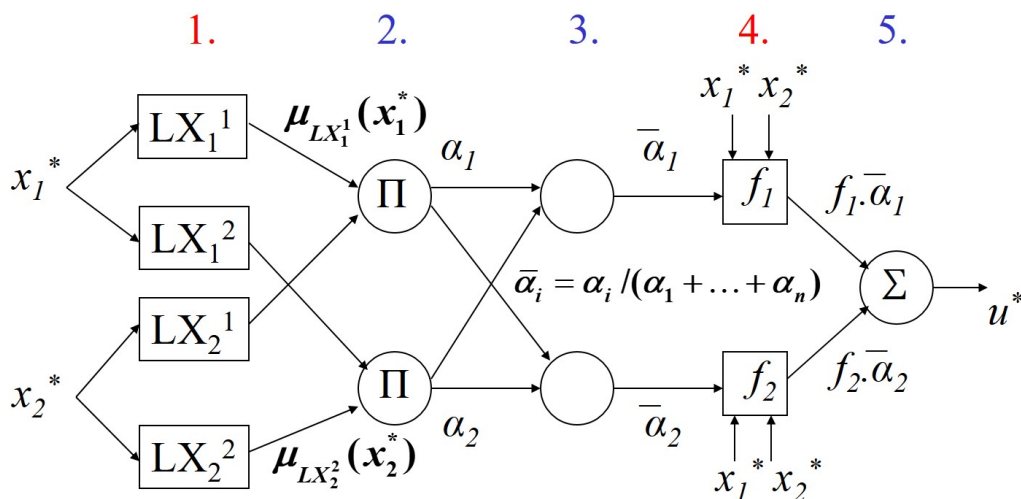
## II. POPIS ALGORITMICKÝCH A HARDVÉROVÝCH PROSTRIEDKOV

V tejto stati popíšeme základné prostriedky, ktoré sú použité pri návrhu systému na reguláciu klímy v skleníku. Jedná sa jednak o popis princípu použitého algoritmu ako aj o popis jednotlivých hardvérových prvkov systému.

Základom systému ANFIS je neurónová sieť, ktorá je funkčne totožná s tzv. TSK (Takagi–Sugeno–Kang) fuzzy regulátorom [4]. Takýto systém sa používa v prípadoch, keď je problematické systémy modelovať, alebo je popis objektov v systéme nejednoznačný. ANFIS sa používa na mapovanie vstupných charakteristík na vstupné funkcie príslušnosti a funkcie príslušnosti na produkčné pravidlá typu Ak–Potom. Následne pomocou týchto pravidiel dostaneme výstupné charakteristiky a na ich základe výstupnú hodnotu v číselnej podobe. Na rozdiel od Mamdaniho fuzzy regulátora tu odpadá nutnosť výpočtovo náročnej defuzzifikácie. Základná architektúra modelu ANFIS je uvedená na Obr. 1.

Pre naše riešenie bolo potrebné vytvoriť taktiež vlastný dátový súbor. Hlavne z dôvodu, že dáta pre takéto riešenia nie sú vo väčšine prípadov všeobecné, keďže nie je všade rovnaké podnebie. Za účelom vytvorenia čo najpresnejšieho dátového súboru sme vytvorili aj model skleníka, do ktorého boli osadené senzory, aby sme získali obraz podmienok pri nasadení v riadnom merítku. Tento dátový súbor pozostáva z hodnôt týchto veličín:

- 1) aktuálna teplota,
- 2) aktuálna vlhkosť,
- 3) odchýlka aktuálnej teploty od požadovanej,
- 4) odchýlka aktuálnej vlhkosti od požadovanej,
- 5) stav ventilátora (stav zapnutý – 100 a stav vypnutý – 0),



Obr. 1 Architektúra neurónovej siete ANFIS.

6) stav zvlhčovača vzduchu (stav zapnutý – 100 a stav vypnutý – 0).

Tieto veličiny sú zaznamenávané každú hodinu. Hodnoty požadovaných veličín závisia od plodiny, ktorú pestujeme, ale väčšinou sú to hodnoty 24°C pre teplotu a 80% pre vlhkosť. Tieto podmienky sú ideálne pre správny rast plodín a zabráneniu plesniam [5].

Raspberry Pi(RPi) je jednodoskový počítač postavený na architektúre typu ARM. Umožňuje spustenie vlastnej verzie Linuxu a v konečnom dôsledku poskytuje takmer všetky možnosti plnohodnotného PC. Veľkými výhodami sú jeho malé rozmery a nízka energetická náročnosť [6]. Toto zariadenie sa vyznačuje svojou jednoduchosťou, nízkou spotrebou energie, či cenou a dostatočným výpočtovým výkonom, ktoré môže byť jednoducho napájané aj energiou zo solárnych panelov, či batérie. Zariadenie využívame na prevádzkovanie webového rozhrania, spojenie s Arduinoom Mega na zber dát a samotnú prácu s modelom ANFIS.

Arduino Mega je platforma založená na mikroregulátore ATmega. Arduino poskytuje programovateľné vstupno/výstupné piny s možnosťou PWM (modulácia pomocou šírky pulzu). Je ho možné napájať z USB, a to nám v tomto prípade ponúka RPi. Zariadenie je využívané ako A/D prevodník pre senzory teploty a vlhkosti a ako regulátor akčných členov [7]. Zariadenie komunikuje s RPi pomocou sériovej linky, kde sú merané hodnoty posielané a taktiež sú aj prijímané hodnoty pre akčné členy.

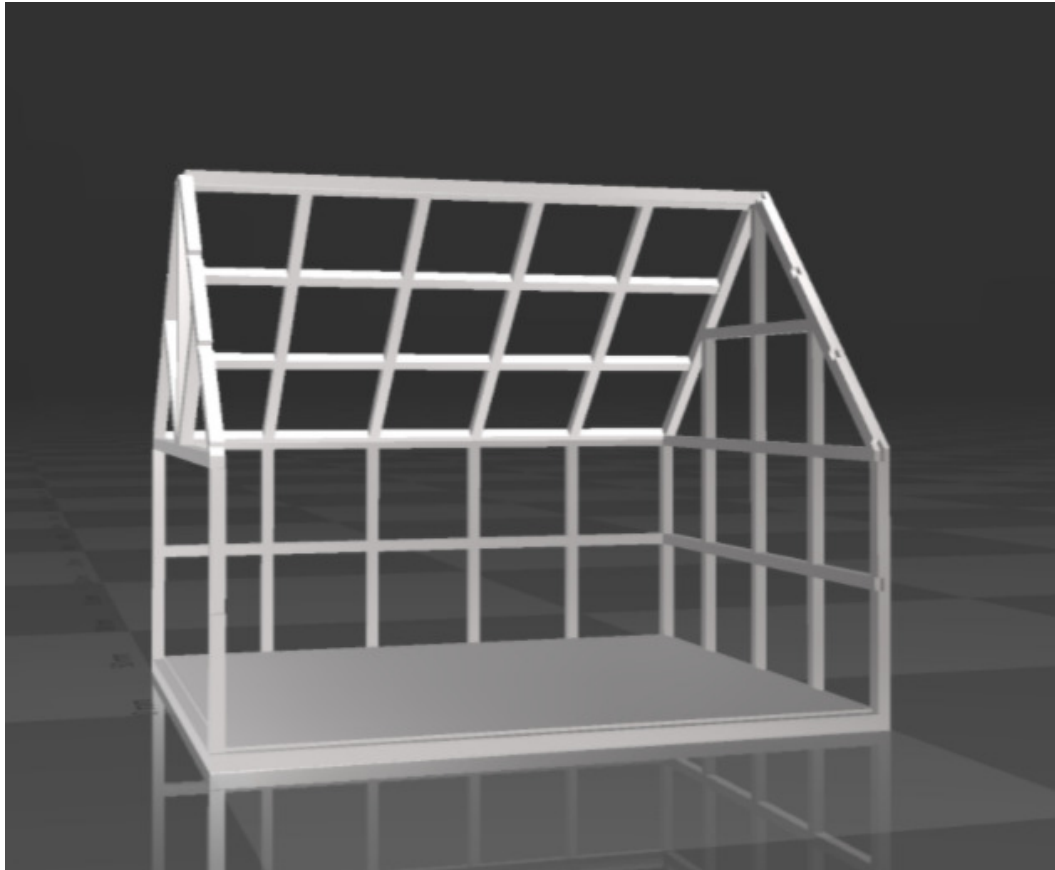
### III. NÁVRH RIEŠENIA

V tejto časti popíšeme implementáciu všetkých zvolených prostriedkov pri návrhu systému na reguláciu klímy v skleníku, a to od tvorby samotného modelu až po implementáciu programovej časti.

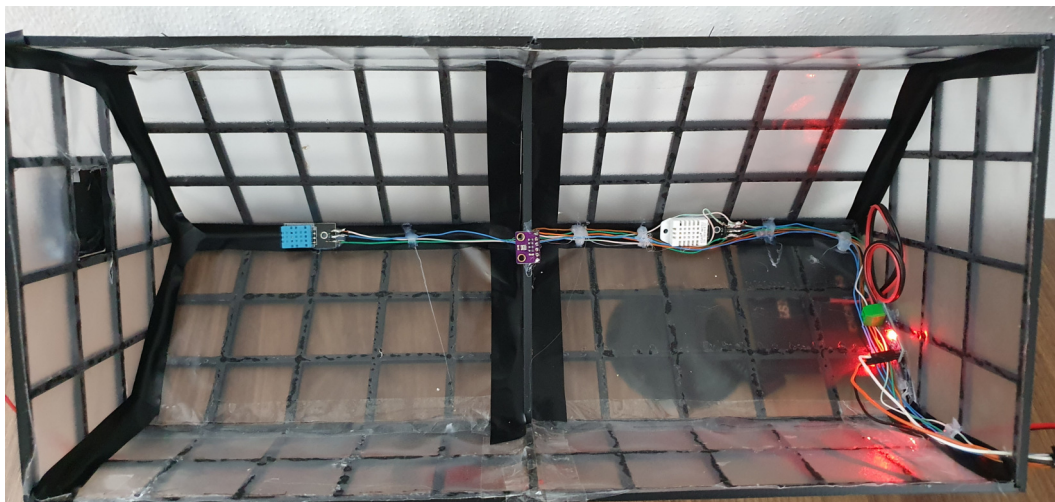
Samotný model skleníka sme zrealizovali pomocou 3D modelovania návrhu a následnej tlači na 3D tlačiarňi. Model má veľkosť 40x20cm, avšak je koncipovaný modulárne, čiže k modelu sa dajú jednoducho pridať ďalšie časti podľa požadovanej veľkosti, vid' Obr. 2.

Po zložení modelu sme do neho osadili spomenuté senzory teploty a vlhkosti DHT 11, DHT 21 a BME 280. Použili sme vyšší počet kusov z dôvodu redundancie, ak by niektorý zo senzorov stratil spojenie alebo nameral nesprávnu hodnotu. Taktiež sme pre budúce experimenty osadili na skleník malý solárny panel o veľkosti 90x70mm výkone 5V/160mA. K nemu je ešte pripojený senzor prúdu ACS712. Toto zapojenie je primárne určené na meranie solárnej radiácie. Následne sa zapojili ešte akčné členy. Použili sme 12V ventilátor o veľkosti 4x4cm s regulovaním otáčok a generický zvlhčovač vzduchu, ktorý funguje na princípe generovania vysokej frekvencie zvuku na kovovom pliešku a následne po kontakte s vlhkou handričkou vytvára vodnú paru, vid' Obr. 3.

Webové rozhranie sme implementovali pomocou nadstavby Flask pre programovací jazyk Python. Umožňuje pomerne jednoduché zostavenie webového rozhrania s možnosťou spúšťania externých Pythonovských scriptov. To sme využili pri načítavaní údajov z SQL databázy. Následne sa tieto dáta spracovali a zobrazili v podobe grafov s aktuálnou hodnotou. Taktiež sme zapracovali aj zobrazovanie stavu akčných členov. Webové rozhranie je navrhnuté tak, aby bolo jednoducho škálovateľné, teda pridanie ďalších hodnôt alebo akčných členov je možné bez zdĺhavej rekonfigurácie zariadenia.



Obr. 2 3D model skleníka.

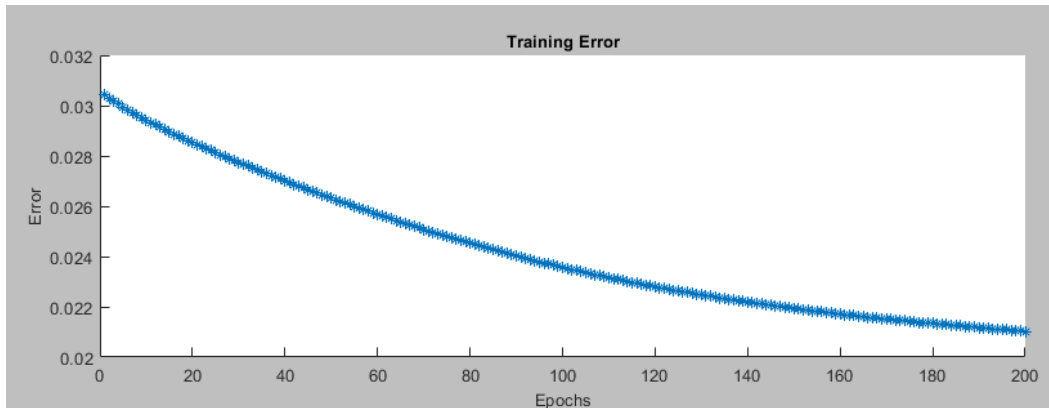


Obr. 3 Hardvérové zapojenie skleníka

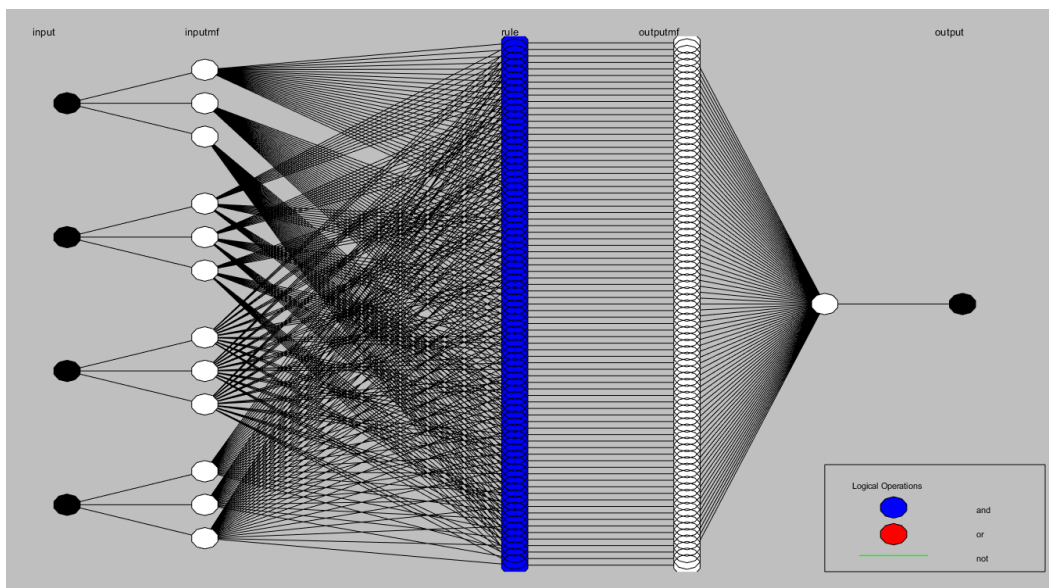
Celkový postup činnosti systému na reguláciu klímy v skleníku je možné opísať nasledovne. Dáta zo senzorov sú spracované pomocou zariadenia Arduino Mega a vyslané na sériovú linku zariadenia. Následne sú z tejto sériovej linky dáta prečítané pomocou RPi, ktoré ich syntakticky rozloží, priloží aktuálnu časovú značku a uloží do SQL databázy. Tento proces prebieha spravidla každú hodinu, avšak jeho interval vyhodnocovania je možné podľa potreby zmeniť.

Model ANFIS bol implementovaný v prostredí jazyka Python s využitím knižnice PyTorch [8]. Pre model sme zvolili štyri vstupné parametre: teplotu, vlhkosť, rozdiel teploty od požadovanej hodnoty a rozdiel vlhkosti od požadovanej hodnoty. Z týchto hodnôt generujeme na výstupe hodnotu v rozmedzí [0;100], kde hodnota 0 predstavuje vypnutý akčný člen a hodnota 100 predstavuje akčný člen spustený na plný výkon. Tvorba modelu v spomínanom prostredí prebieha

v niekoľkých krokoch, a to načítaní dát, nastavení funkcií príslušností, trénovaní a testovaní modelu. Použili sme 120 záznamov, čo predstavuje meranie hodnôt za posledných päť dní. Tieto dáta sme si rozdelili v pomere 70%:30% na trénovacie a testovacie. Následne sme si zvolili lichobežníkové funkcie príslušnosti, ktoré najviac vyhovovali našim dátam. Každéj vstupnej hodnote sme nastavili tri vstupné funkcie príslušnosti. Po vygenerovaní pravidiel a rozsahov pre jednotlivé stupne príslušnosti sme spustili trénovanie na 200 epochách, kde model dosiahol najnižšiu chybovú funkciu, vid' Obr. 4. Výsledkom je štruktúra pravidiel, ako ju vidieť na Obr. 5.



Obr. 4 Vývoj chyby modelu ANFIS počas trénovania v jednotlivých epochách.



Obr. 5 Štruktúra pravidiel modelu ANFIS.

Výsledná trénovacia chyba vyhodnotená metrikou MSE(mean square error) bola 0,02 a výsledná presnosť cca 74%. Model obsahuje 81 produkčných pravidiel. Model sme následne vedeli vyhodnotiť na aktuálnych dátach. Takéto modely ANFIS sme potrebovali vytvoriť dva, jeden pre hodnotu ventilácie a jeden pre hodnotu zvlhčovania, keďže nebolo možné zahrnúť výkon oboch akčných členov do jedinej hodnoty.

Po nasadení navrhnutého zariadenia do reálnej prevádzky sa v hodinových intervaloch vykonávali merania, na základe ktorých zariadenie riadilo prevádzku ventilácie a zvlhčovania. Tabuľka 1 ukazuje postupnosť riadenia ventilácie a zvlhčovania v závislosti od vstupných hodnôt. Ako je vidieť, je možné badať isté ovplyvňovanie vstupných veličín na obidva výstupy, takže systém negeneruje rovnaké výstupy iba v rámci jednej veličiny, ale vplýva naň aj druhá vstupná veličina (napr. na zvlhčovanie nevlýva iba samotná úroveň vlhkosti ale aj teplota), čo nepovažujeme za negatívny jav, nakoľko v skutočnosti teplota a vlhkosť sa navzájom ovplyvňujú.



Tabuľka 1

Príklad riadenia ventilácie a zvlhčovania pomocou modelu skleníka s využitím siete typu ANFIS.

Teplota	Vlhkosť	Teplotná odchýlka	Odchýlka vlhkosti	Ventilátor	Zvlhčovač
24	80	0	0	0,002	0,001
26	85	2	5	45,598	39,861
28	95	4	15	99,7039	95,420
27	90	3	10	66,357	70,583
26	80	2	0	43,216	0,948

#### IV. ZÁVER

V konečnom dôsledku hodnotíme výstup tohto riešenia kladne, keďže sa nám podarilo vytvoriť funkčnú vlastnú škálovateľnú hardvérovú a softvérovú platformu na testovanie riešení zameraných na ovládanie veličín v skleníku. Taktiež sa nám podarilo zostaviť si vlastný dátový súbor a natréňovať na ňom dostatočne použiteľný model ANFIS. Po odstránení nedostatkov, hlavne zmeny platformy ARM na x86, bude toto riešenie možné bez problémov nasadiť aj na reálny skleník, kde by sa dátový súbor tvoril ešte jednoduchšie a s oveľa presnejšími dátami.

#### LITERATÚRA

- [1] B. Khoshnevisan, S. Rafiee, M. Omid, and H. Mousazadeh, "Development of an intelligent system based on ANFIS for predicting wheat grain yield on the basis of energy inputs," *Information Processing in Agriculture*, vol. 1, no. 1, pp. 14–22, 2014.
- [2] I. Lopez-Cruz and L. Hernández-Larragoiti, "Neuro-fuzzy models for air temperature and humidity of a greenhouse," *Acta Horticulturae*, vol. 927, pp. 611–617, 02 2012.
- [3] A. Karpenko and I. Petrova, "The conceptual model of neuro-fuzzy regulation of the microclimate in the room," in *18th IFAC Conference on Technology, Culture and International Stability TECIS 2018*, vol. 51, no. 30, 2018, pp. 636–640.
- [4] D. Driankov, H. Hellendoorn, and M. Reinfrank, *An introduction to fuzzy control*. Springer, 1996.
- [5] Drygair, "What are the optimal humidity and temperature set points for greenhouse growing?" [online] <https://drygair.com/what-are-the-optimal-humidity-and-temperature-set-points-for-greenhouse-growing/>, 2019.
- [6] "Teach, learn, and make with raspberry pi," [online] <https://www.raspberrypi.org/>, 2021.
- [7] "Arduino home," [online] <https://www.arduino.cc/>, 2021.
- [8] "Pytorch," [online] <https://pytorch.org/>, 2021.

# Nástroje a knižnice pre analýzu dát v oblasti medicíny

<sup>1</sup>Daniel NOVÁK, <sup>2</sup>Michal KOLÁRIK

Katedra kybernetika a umelej inteligencie, Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach, Slovenská Republika

<sup>1</sup>daniel.novak@student.tuke.sk, <sup>2</sup>michal.kolarik@tuke.sk

**Abstract** – V tomto príspevku sme sa venovali nástrojom a knižniciam, ktoré sú určené pre spracovanie medicínskych dát. Analyzovali sme komerčné ale aj voľne dostupné nástroje a knižnice používané pri analýze medicínskych dát. Následne sme vybrali jednu knižnicu na ktorej sme vykonávali experimenty pre jednotlivé fázy CRISP-DM modelu na troch odlišných datasetoch. Knižnicu sme vyhodnocovali porovnávaním s bežne zaužívaným postupom analýzy dát pre jazyk R.

**Keywords** – dátová analýza, Healthcare.ai, medicínske dáta

## I. ÚVOD

V súčasnej dobe algoritmy strojového učenia pomáhajú ľuďom v rôznych oblastiach života. Od zábavného priemyslu, cez marketing a personalizované reklamy až k priemyslu a výrobe. Dôležitou oblasťou je aj oblasť medicíny, kde strojové učenie môže pomôcť od šetrenia nákladov nemocniciam, lepšej identifikácii chorôb až po záchranu ľudských životov.

Analýza dát v oblasti medicíny má častokrát svoje špecifiká, čo vyžaduje niekedy odlišný prístup a použitie špeciálnych nástrojov a prístupov. Okrem špecifického typu dát, kde je potrebná účasť lekára pri dátovej analýze, je potrebné myslieť aj na bezpečnosť dát, keďže častokrát ide o citlivé dáta o zdravotnom stave pacientov. Objem dát je ďalším problémom. Zatiaľ čo niekedy býva obrovské množstvo dát, ktoré je problematické spracovať, stretávame sa aj s opačným problémom a to je nedostatok zozbieraných dát z ktorých by sme potrebovali získať dôležité znalosti. Pre riešenie týchto problémov sú nevyhnutné správne postupy a využitie nástrojov, ktoré uľahčujú spracovanie dát.

## II. ANALÝZA SÚČASNÉHO STAVU

### A. Komerčné nástroje

Pri dátovej analýze v medicínskej oblasti sme skúmali použitie komerčných nástrojov. Pri väčšine týchto nástrojov sa použitie v oblasti medicíny nelíši od iných oblastí a moduly, ktoré slúžia pre spracovanie medicínskych dát boli už priamo implementované v nástroji.

Medzi populárne komerčné nástroje využívané v oblasti medicíny patria:

- Stata [1] - komerčný štatistický softvér, ktorý poskytuje nástroje pre analýzu tabuľkových dát, správu dát a grafické výstupy v podobe rôznych grafov.
- SAS [2] - štatistický softvérový balík využívaný najmä pre štatistickú analýzu a vizualizáciu údajov. Poskytuje nástroje pre pokročilú analýzu, hĺbkovú analýzu, podporu rôznych formátov a databázových systémov.
- IBM SPSS [3] - softvérový balík pozostávajúci z troch nástrojov SPSS Statistics - zameraný na ľahko použiteľné užívateľské prostredie pri vytváraní štatistík, SPSS Modeler - jednoduchý vizuálny nástroj na tvorbu analytických modelov a SPSS Amos ktorý ponúka pokročilé štatistické funkcie a testy.
- Matlab [4] - komerčný nástroj pôvodne určený pre efektívnu a rýchlu prácu s maticami a pokročilými matematickými operáciami, ktorý bol postupne doplnený o funkcie pre analýzu tabuľkových dát.
- Microsoft Excel [5] - jeden z najpoužívanejších tabuľkových softvérov určený pre vytváranie, editáciu a správu tabuliek. Ponúka funkcie pre jednoduchú filtráciu, triedenie, analýzu dát a vytváranie prehľadných tabuliek a grafických výstupov.

Table 1  
 Porovnanie dostupných knižníc pre analýzu medicínskych dát

Nástroj	Healthcare.ai	SCALPEL3	Pumas
Podporované jazyky	R/Python	Scala/Python	Julia
Opensource	áno	áno	nie
Využitie	dolovanie dát a dátová analýza	štatistická analýza komplexných dát v medicíne	podpora a analýza pri vývoji liekov
Objem spracovaných dát	stredne veľké datasety	veľmi veľké a komplexné datasety	stredne veľké datasety
Dokumentácia	silná	slabá	dobrá
Komunita	dobrá	žiadna	dobrá
1. fáza CRISP-DM pochopenie cieľa	nie	nie	nie
2. fáza CRISP-DM pochopenie dát	áno	áno	áno
3. fáza CRISP-DM príprava dát	áno	áno	áno
4. fáza CRISP-DM modelovanie	áno	áno	áno
5. fáza CRISP-DM vyhodnotenie výsledkov	áno	áno	áno
6. fáza CRISP-DM nasadenie	nie	nie	nie

### B. Opensource nástroje a knižnice

Opensource nástroje využívajú otvorený zdrojový kód do ktorého môže ktokoľvek prispievať a vylepšiť tak ich možnosti. Zvyčajne sú tieto nástroje vytvorené v programovacom jazyku, ktorý má tiež otvorený zdrojový kód a je voľne dostupný. Najčastejšími jazykmi v oblasti dátovej analýzy sú:

- R - voľne dostupný, štatistický jazyk, využívaný najmä v akademickej a vedeckej sfére. Ponúka rýchlu prácu s tabuľkami, zoznamami, vektormi. Umožňuje vykonávať štatistické výpočty a robí grafické vizualizácie nad dátami.
- Python - vysokoúrovňový, objektovo orientovaný jazyk, ktorý má širokú škálu aplikovateľnosti v oblasti informatiky. Je veľmi populárny pre svoju prehľadnú syntax a efektívnosť písania kódu. V oblasti dátovej analýzy pre python existuje množstvo modulov a knižníc, ktoré uľahčujú spracovanie dát.
- Julia - ide o najmladší z týchto troch jazykov. Bol vytvorený za účelom efektívnosti a rýchlosti spracovania numerických a tabuľkových dát. Má v sebe množstvo matematických a štatistických funkcií ako aj funkcie pre spracovanie veľkých dát.

Nasledujúce knižnice, ktoré sú určené pre analýzu dát v medicínskej oblasti, využívajú niektorý z predchádzajúcich voľne dostupných jazykov.

- Healthcare.ai [6] - knižnica pre dolovanie v medicínskych dátach určená pre jazyky R a Python. Knižnica ponúka funkcie pre zjednodušenie celkovej analýzy dát, ako napríklad strojové učenie, profilovanie dát, príprava dát, predikcia, klasifikácia.
- Pumas [7] - knižnica založená na jazyku Julia, ktorá obsahuje viacero modulov pre umelú inteligenciu a dátovú analýzu najmä pre farmaceutický priemysel. Využíva sa pri predklinických a klinických štúdiách vývoja liekov a ponúka sadu nástrojov pre kvantitatívnu analýzu a efektívne spracovanie dát. Nie je celá opensource ale je dostupná zdarma pre vedecké účely.
- SCALPEL3 [8] - je škálovateľný softvérový rámec pre analýzu veľkých dát z oblasti medicíny. Pozostáva z troch knižníc a to: SCALPEL-Flattening - knižnica zodpovedná za denormalizáciu veľkého množstva dát, SCALPEL-Extraction - knižnica ktorá poskytuje rýchlu extrakciu dát a SCALPEL-Analysys - knižnica, ktorá umožňuje interakciu, manipuláciu a analýzu dát, ktoré boli predspracované predchádzajúcimi SCALPEL knižnicami.

Vyššie opísané knižnice sme navzájom porovnali v tabuľke 1. Na základe porovnania sme sa rozhodli ďalej pracovať s knižnicou Healthcare.ai a zistiť jej prínosy pri spracovaní medicínskych dát.

### III. POUŽITÉ KNIŽNICE A METÓDY

Pracovali sme s knižnicou Healthcare.ai a jej implementáciou v štatistickom jazyku R. Nakoľko knižnica implementuje všetky potrebné funkcie pre analýzu dát, prípravu a čistenie dát ako aj modelovanie dát, stačilo nám pre jej testovanie iba načítanie knižnice *healthcareai*. Knižnica pri

fáze modelovania v základnom nastavení trénuje súčasne 3 modely a to Random Forest, eXtreme Gradient Boosting a Glmnet, pričom automaticky nastaví vhodné hyperparametre a nakoniec porovná výsledky natrénovaných modelov na základe zvolenej metriky.

Keďže sme chceli porovnať možnosti tejto funkcie s bežným postupom dátovej analýzy a modelovania v jazyku R, museli sme si v tomto prípade načítať potrebné knižnice pre predspracovanie dát a tvorbu rovnakých troch modelov, a to: *dplyr*, *naniar*, *randomForest*, *xgboost*, *glmnet* a *caret*

#### A. Použité modely

- Random Forest [9] - súborová metóda ktorá kombinuje veľký počet náhodných stromov, pričom cieľová trieda sa vyhodnotí na základe hlasovania z výsledkov jednotlivých stromov.
- eXtreme Gradient Boosting [10] - podobne ako v prípade modelu Random Forest ide o súbor jednoduchších modelov, zvyčajne stromov, ktoré sú vytvárané po fázach a sú v každej fáze optimalizované na základe stratovej funkcie.
- Glmnet [11] - je to balík jazyka R ktorý zároveň poskytuje rýchle a efektívne funkcie pre lineárne a regresné modely.

### IV. MODELOVANIE POMOCOU METODIKY CRISP-DM

#### A. Pochopenie cieľa

Ako cieľ tejto štúdie sme porovnávali možnosti knižnice Healthcare.ai s možnosťami jazyka R a jeho modulov pri analýze troch vybraných datasetov z oblasti medicíny. Pre objektívne zhodnotenie sme vybrali datasety, ktoré sa odlišovali svojím charakterom či už veľkosťou, počtom chýbajúcich hodnôt alebo cieľovou úlohou. Pre každý z datasetov sme vytvárali modely strojového učenia, zisťovali sme rozdiely postupov a krokov pri použití knižnice Healthcare.ai a bežného prístupu v jazyku R. Počas tohto procesu sme zaznamenávali zložitosť vytvorenia modelu, počet riadkov kódu, zložitosť odstránenia chýbajúcich hodnôt, dostupnosť nápovedy či rýchlosť odozvy.

#### B. Pochopenie dát

V štúdiu sme pracovali s tromi datasetmi. Zdrojom týchto datasetov bola databáza UCI <sup>1</sup>. Vybrané datasety sú z oblasti medicíny a líšia sa veľkosťou dát, počtom atribútov či chýbajúcich hodnôt a cieľovou úlohou.

- Prvý dataset - dataset o rakovine prsníka je získaný z "University of Wisconsin Hospitals" [27]. Dataset obsahuje 699 záznamov a 11 atribútov, ktoré zachytávajú rôzne ukazovatele dôležité pre určovanie rakoviny. Cieľovou úlohou bola binárna klasifikácia.
- Druhý dataset - dataset zaznamenávajúci rečové signály získané telemonitorovacím zariadením počas 6 mesiacov od ľudí, ktorí majú počiatočné štádium parkinsonovej choroby. Tento dataset vytvoril Athanasios Tsanas [28]. Pozostáva z 5875 záznamov a 22 atribútov. Cieľovou úlohou bola regresia.
- Tretí dataset - dataset o ochoreniach štítnej žľazy získaných inštitúciou Garavan a J. Ross Quinlan [29]. Dataset obsahuje 2800 záznamov a 30 atribútov. Cieľovou úlohou bola klasifikácia.

#### C. Príprava dát

V tejto fáze sme pristupovali k zvoleným datasetom jednotlivo, pričom sme skúmali rozdiely pri použití knižnice Healthcare.ai a štandardného prístupu a knižnic jazyka R.

Prvý dataset obsahoval chýbajúce dáta, ktoré boli označené znakom "?" a s ktorými sme sa potrebovali vysporiadať. Knižnica Healthcare.ai dokáže pracovať s chýbajúcimi hodnotami, ktoré ale musia byť v správnom tvare. Knižnica nám preto automaticky ponúkla nápovedu ako zmeniť hodnoty "?" na hodnoty "NA". Okrem toho táto knižnica zisťuje cieľovú úlohu na základe určeného cieľového atribútu, preto bolo potrebné zmeniť číselné hodnoty cieľového atribútu za kategorické. Knižnica pri modelovaní ponúka aj možnosť vynechať vybrané atribúty, preto sme nemuseli v tomto kroku odstrániť z tabuľky atribút *Sample code number*. V prípade použitia štandardných knižnic jazyka R bolo potrebné vykonať rovnaké úkony s chýbajúcimi dátami a zmenu cieľového atribútu na kategorický. Navyše bolo nutné odstrániť atribút *Sample code number* pred modelovaním. Okrem toho bolo nutné pretransformovať všetky dáta okrem cieľového atribútu na numerické. Pre model eXtreme Gradient Boosting sme museli zmeniť aj cieľový atribút na numerický.

Druhý dataset vyžadoval menšie množstvo operácií. V prípade knižnice Healthcare.ai išlo iba o odstránenie prvého riadku, ktorý obsahoval chybné načítané dáta. Naopak v prípade použitia

<sup>1</sup><https://archive.ics.uci.edu/>

štandardných knižníc jazyka R, bolo nutné okrem odstránenia prvého riadku odstrániť aj atribút *subject#* ktorý neposkytoval relevantné informácie. Pred modelovaním dát bola nutná ešte úprava názvov atribútov, pretože špeciálne znaky spôsobovali niektorým použitým modelom problémy.

V treťom datasete boli v oboch skúmaných prípadoch nahradené hodnoty "?" za hodnoty "NA". Pri použití štandardných knižníc bolo nutné ešte upraviť cieľový atribút pre jednotlivé modely rovnako ako v prvom datasete.

#### D. Modelovanie

V tejto fáze sme sa venovali vytváraniu modelov na predspracovaných dátach. Dataset sme rozdelili so zachovaním distribúcie cieľovej triedy v pomere 80:20 na tréningovú a testovaciu množinu.

Knižnica *Healthcare.ai* ponúka automatickú funkciu, ktorá na základe vstupných dát, cieľového atribútu a metriky natrénuje viaceré modely, ktoré následne porovná. V našom prípade šlo o modely *Random forest*, *eXtreme Gradient Boosting* a model *Glmnet*. Tieto tri modely sme vytvárali aj použitím štandardných knižníc jazyka R.

V prípade modelu *Random forest* bolo nutné načítať knižnicu *naniar* a *randomForest*. Tento model potrebuje na vstupe numerické dáta, avšak dokáže pracovať aj s "NA" hodnotami.

V prípade modelu *eXtreme Gradient Boosting* bolo nutné načítať knižnicu *xgboost* a *dplyr* a *naniar*. Tento model si nedokáže poradiť s "NA" hodnotami, preto boli v našom prípade takéto záznamy odstránené. Model pracuje iba s numerickými hodnotami a cieľový atribút musí byť rovnako numerický.

Pri vytváraní modelu pomocou balíčka *glmnet* bolo nutné načítať knižnicu *glmnet* a *naniar* a *caret*. Tento model si vie poradiť s chýbajúcimi hodnotami. Pre klasifikačnú úlohu bolo nutné zmeniť cieľový atribút na typ faktor.

#### E. Vyhodnotenie

Na základe experimentov na troch rozdielnych datasetoch a jednotlivých krokoch predspracovania a tvorby modelov môžeme povedať, že použitie knižnice *Healthcare.ai* oproti bežnému postupu v jazyku R, zjednodušuje celý proces analýzy dát. Už pri predspracovaní dát nám okrem jednoduchšej identifikácie a odstránenie chýbajúcich hodnôt ponúkala *Healthcare.ai* knižnica aj nápovedu v podobe ďalších krokov. Pozitívne hodnotíme aj dostatočne dobrú dokumentáciu tejto knižnice.

Pri použití bežného prístupu sme pre vytvorenie modelov potrebovali načítať v priemere 4 rôzne knižnice, pričom pri knižnici *Healthcare.ai* to bola iba táto jedna knižnica. Pre celý proces analýzy dát sme pri bežnom prístupe potrebovali v priemere dvojnásobný počet riadkov zdrojového kódu ako pri použití *Healthcare.ai* knižnice. Okrem toho knižnica *Healthcare.ai* ponúkala nápovedu pri modelovaní pre zmenu typu cieľového atribútu a teda aj cieľovej úlohy (klasifikácia / regresia) ako aj rôzne varovania a hlásenia chýb. Tieto funkcie nám pri ostatných knižniciach a bežnom prístupe v jazyku R chýbali.

## V. ZÁVER

V tomto článku sme sa venovali analýze existujúcich nástrojov a knižníc používaných pri spracovaní medicínskych dát. Na základe porovnania dostupných knižníc sme zvolili najvhodnejšiu knižnicu s ktorou sme ďalej pracovali a skúmali sme jej prínos pri spracovaní medicínskych dát. Zistili sme, že použitie takejto knižnice môže veľmi zjednodušiť a zrýchliť proces analýzy dát od fázy predspracovania cez modelovanie a vyhodnotenie. Knižnica ponúkala nápovedu k ďalšiemu postupu pri spracovaní dát. Rovnaké výsledky je možné dosiahnuť aj použitím bežných postupov avšak s väčším úsilím pri pochopení dát, príprave dát, ako aj pri vytváraní a ladení modelov.

## VI. POĎAKOVANIE

Táto práca vznikla vďaka podpore Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-16-0213 a Vedeckej grantovej agentúry MŠVVaŠ SR a SAV projekt č. 1/0685/21.

## REFERENCES

- [1] S. LLC. (2021) Stata Statistical Software: release 17. college station. [Online]. Available: <https://www.stata.com/features/>
- [2] SAS. (2021) What can I do with SAS? [Online]. Available: [https://support.sas.com/software/products/sas-studio/faq/SAS\\_whatism.htm](https://support.sas.com/software/products/sas-studio/faq/SAS_whatism.htm)
- [3] M. N. (2019) Your ultimate guide to SPSS Statistics vs SPSS Modeler. [Online]. Available: [https://community.ibm.com/community/user/datascience/blogs/nitin-mathur1/2019/11/14/spss-statistics-vs-modeler?\\_ga=2.78907476.1127188112.1605213120-228165455.1603795466](https://community.ibm.com/community/user/datascience/blogs/nitin-mathur1/2019/11/14/spss-statistics-vs-modeler?_ga=2.78907476.1127188112.1605213120-228165455.1603795466)

- [4] T. M. Inc. (2021) MATLAB the language of technical computing. [Online]. Available: [https://www.mathworks.com/help/matlab/index.html?s\\_tid=CRUX\\_lftnav](https://www.mathworks.com/help/matlab/index.html?s_tid=CRUX_lftnav)
- [5] Microsoft Corporation, "Microsoft excel." [Online]. Available: <https://office.microsoft.com/excel>
- [6] Health Catalyst, "Healthcareai." [Online]. Available: <https://healthcare.ai>
- [7] Pumas, "Version 1.1.0." [Online]. Available: <https://docs.pumas.ai/stable/>
- [8] E. Bacry, S. Gaïffas, F. Leroy, M. Morel, D.-P. Nguyen, Y. Sebiat, and D. Sun, "Scalpel3: a scalable open-source library for healthcare claims databases," *International Journal of Medical Informatics*, p. 104203, 2020.
- [9] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002. [Online]. Available: <https://CRAN.R-project.org/doc/Rnews/>
- [10] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 785–794. [Online]. Available: <https://doi.org/10.1145/2939672.2939785>
- [11] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for cox's proportional hazards model via coordinate descent," *Journal of Statistical Software*, vol. 39, no. 5, pp. 1–13, 2011. [Online]. Available: <https://www.jstatsoft.org/v39/i05/>

# Simulovanie LTE a 5G prístupových technológií z pohľadu energetickej účinnosti s využitím NS-3

<sup>1</sup>Maroš BAUMGARTNER, <sup>2</sup>Jozef JUHÁR, <sup>3</sup>Ján PAPAĽ

<sup>1,2,3</sup> Katedra elektroniky a multimediálnych telekomunikácií,  
Fakulta elektrotechniky a informatiky,  
Technická univerzita v Košiciach,  
Slovenská republika

<sup>1</sup>maros.baumgartner@tuke.sk, <sup>2</sup>jozef.juhar@tuke.sk, <sup>3</sup>jan.papaj@tuke.sk

**Abstrakt** — V dnešnej dobe sa zvyšujú požiadavky na prenos dát a efektívnosť sietí IoT. Efektivitu siete na všetkých úrovniach možno zvýšiť použitím sietí 5G. V tomto článku simulujeme a analyzujeme LTE, enhanced Mobile BroadBand a enhanced Mobile BroadBand so službami milimetrových vln v scenároch s rôznym počtom uzlov IoT a analyzujeme dosiahnuté výsledky v oblasti spotreby energie a energetickej účinnosti. Simulované scenáre a výsledky sa získali pomocou NS-3. Tento článok popisuje použitý rámec mmWave a energy model pre NS-3, pretože boli hlavnou časťou v tejto a budúcej práci.

**Kľúčové slová** — mmWave, eMBB, IoT, 5G, NS-3, simulation

## I. ÚVOD

Siete novej generácie, ktoré sú známe ako 5G siete sa stávajú stále väčšou súčasťou nášho života. Tieto siete by mali priniesť vylepšenia v rôznych oblastiach prenosu dát, napríklad kapacita, end – to – end oneskorenie, robustnosť, škálovateľnosť, rýchlosť prenosu dát a energetická účinnosť.

5G umožňuje rýchlejšiu, stabilnejšiu a bezpečnejšiu konektivitu, ktorá v sebe zahŕňa samo riadiace vozidlá, inteligentné siete pre obnoviteľnú energiu a roboty, ktoré podporujú umelú inteligenciu. Všetky tieto statické a dynamické snímače získavajú veľké množstvo IoT dát, ktoré sa používajú na komunikáciu a zlepšenie efektívnosti siete.

Jednými z najdôležitejších parametrov v 5G sieťach a IoT (Internet of Things) sú energetická účinnosť a spotreba energie jednotlivých IoT zariadení v sieti. Hlavným cieľom energetickej účinnosti siete je okrem zníženia nákladov na prevádzku siete hlavne zníženie spotreby energie a predĺženie životnosti batérií v koncových IoT zariadeniach, napríklad mobilné telefóny, laptopy, drony a podobne.

Kvôli tomu, že v súčasnej dobe je LTE (Long Term Evolution) stále najrozšírenejšia služba na prenos dát v sieti sme sa v tomto článku zamerali na porovnanie dosiahnutých výsledkov energetickej účinnosti a spotreby energie starších LTE sietí s technológiami 5G sietí v podobe enhanced Mobile BroadBand a milimetrových vln.

Článok je rozdelený do niekoľkých častí. V prvej časti opisujeme implementovaný mmWave (millimeter wave) model v NS-3 (Network Simulator 3), ktorý bol použitý pri simuláciách. Táto časť tiež opisuje použitú technológiu eMBB (enhanced Mobile BroadBand). Okrem mmWave modelu opisuje aj energetický model pomocou ktorého sme získali a analyzovali dosiahnuté výsledky.

V nasledujúcej časti sú popísané simulačné scenáre, použité parametre a dosiahnuté výsledky. Boli porovnané priemerné hodnoty energetickej účinnosti a spotrebovanej energie IoT uzlov LTE služby s výsledkami získanými simuláciou 5G sietí.

## II. NS-3 RÁMCE

V tejto časti je opísaný mmWave a energetický model pre simulátor NS-3.

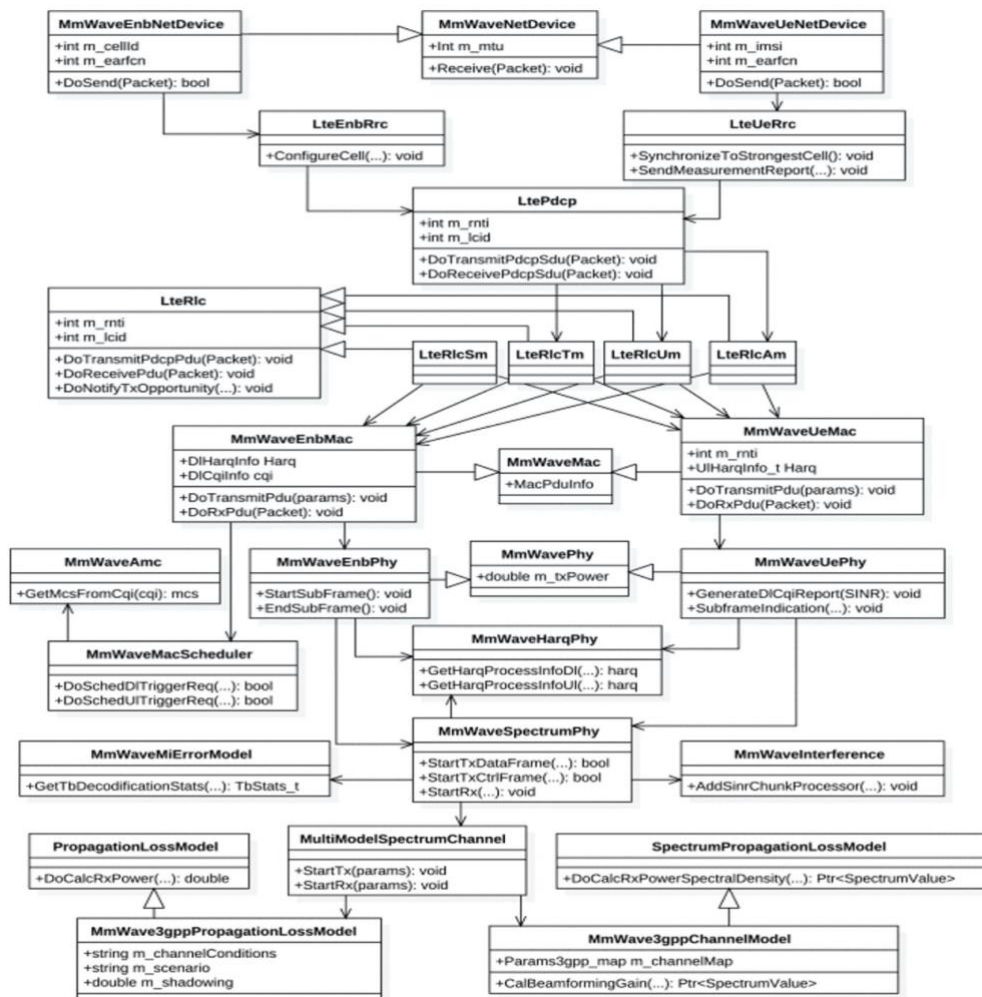
### A. mmWave model

Základom modulu mmWave pre NS-3 je model LENA. Je považovaný za najrobustnejší. Bol navrhnutý tak, aby simuloval komplexné mobilné siete v štýle 3GPP (3rd Generation Partnership Project) [1]. Okrem protokolov LTE / EPC implementuje aj svoje vlastné vrstvy MAC a PHY [2][3]. Tieto vrstvy podrobne popísali Rebato a Mezzavilla vo svojej práci [4][5].

Základný diagram UML modulu mmWave, ktorý popisuje vzťahy medzi triedami a vrstvami, je znázornený na Obr. 1. Celý diagram predstavuje štruktúru end-to-end simulátora.

Triedy *MmWaveEndNetDevice* a *MmWaveUeNetDevice* majú funkciu rádiových zásobníkov pre mmWave eNodeB a mmWave UE. Trieda *McUeNetDevice* navyše zaisťuje, že sa zariadenie *NetDevice* môže pripojiť k technológiám mmWave a LTE pomocou duálneho zásobníku.

Triedy vrstiev MAC, *MmWaveEnbMac* a *MmWaveUeMac* implementujú SAP (Service Access Point), užívateľské rozhranie a LTE modul pre spoluprácu s LTE RLC vrstvou [6]. MAC, *MmWaveMacScheduler* a odvodené triedy implementujú podporu pre triedy zo skupiny RLC, t. j. TM (transparentný režim), UM (nepotvrdený režim), SM (režim sýtosti) a AM (režim potvrdenia). SAP je implementovaný v triede *LteEnbRrc* pomocou plánovača MAC z dôvodu konfigurácií na LTE vrstve riadenia rádiových zdrojov (RRC) [6][7].



Obr. 1 UML diagram mmWave modulu simulátora NS-3

Triedy zo skupiny *MmWavePhy* poskytujú smerový prenos, príjem dát cez downlink a uplink. Starajú sa tiež o kanály, ktoré fungujú na princípe riadenia MAC správ [1]-[3]. Inštancie triedy *MmWaveSpectrumPhy* sa používajú na komunikáciu inštancií triedy *Phy* prostredníctvom *SpectrumChannel* [7]. Trieda *MmWaveSpectrumPhy* a jej inštancie sú zdieľané pre uplink a downlink vďaka fyzickej vrstve modulu mmWave, ktorá je založená na TDD (Time Division Duplexing) [8][9].

### B. Energetický rámec

Spotreba energie je kľúčová vlastnosť pre bezdrôtové IoT uzly v mobilných sieťach. Pre účely výskumu o spotrebe energie bol do simulátora NS-3 implementovaný energy framework



pomocou ktorého boli získané dáta pre ďalšie spracovanie [10]. Energetický model NS-3 sa skladá z 3 častí:

- Energy source
- Energy model
- Energy harvester

Nižšie je opísaný zdroj energie a energetický model z dôvodu využitia pri simulovaní.

#### 1) Energy source

Energy source v sieti reprezentuje zdroj energie v každom IoT uzle. Každý uzol môže byť pripojený do viacerých modelov device energy. V prípade, že bude pripojený zdroj energie do takýchto modelov bude konkrétne zariadenie získavať energiu z tohto zdroja [10][11].

Základnou funkciou zdroja energie je poskytovať energiu pre uzol. Ak sa energia v uzly vyčerpá daný uzol informuje susedné uzly, ktoré môžu reagovať na túto udalosť, získať informácie o zvyšnej energii alebo o úrovni nabitia batérie.

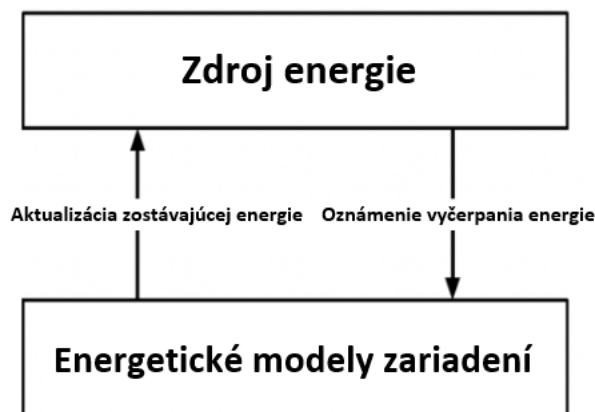
Aby bolo možné simulovať napájacie zdroje musí trieda energy source zachytiť 2 dôležité efekty praktických batérií:

- Efekt frekvenčnej kapacity - Zníženie životnosti batérie, keď je prúdový odber vyšší ako menovitá hodnota batérie.
- Efekt obnovy - Predĺženie životnosti batérie, keď sa batéria striedavo prepína medzi vybitým a pokojovým stavom [10].

Trieda energy source rozdeľuje uzol na viac menších zariadení. Každé z týchto zariadení spotrebuje energiu samostatne. Energy source sa periodicky pýta na spotrebovanú energiu zo všetkých zariadení v tom istom uzle z ktorej vypočíta celkovú spotrebu energie.

#### 2) Energy model

Model pre spotrebu energie v uzloch IoT je energetická trieda zariadenia. Každé zariadenie má definovaných niekoľko stavov a každý stav je spojený s určitou hodnotou spotreby energie. Ak sa stav zariadenia zmení, energetický model upozorní energetickú triedu na rozdiel v aktuálnej spotrebe energie zariadenia. Na základe toho zdroj energie vypočíta aktuálnu spotrebu energie a aktualizuje zostávajúcu energetickú hodnotu. Obr. 2 zobrazuje blokovú schému energetického modelu v NS-3 [10][11].



Obr. 2 Blokový diagram energetického modelu v simulátore NS-3 [12]

### III. SIMULAČNÉ PARAMETRE A SCENÁRE

Táto časť obsahuje popis použitých parametrov simulácií, simulačných scenárov a implementované metódy v simulátore NS-3.

Táto práca sa zameriava na simulovanie a porovnanie jednotlivých metód prenosu dát pre výskum v oblasti robustného prenosu dát pomocou 5G sietí. Práca porovnáva a analyzuje výsledky simulácií LTE, eMBB a eMBB s mmWave prenosových metód z pohľadu energetickej účinnosti a spotrebovanej energie siete pri rôznych typoch scenárov, ktoré boli vytvorené pomocou NS-3. Tento typ novej služby 5G bol zvolený najmä preto, lebo má prísny a dobre definovaný štandard.

#### A. Simulačné parametre

V tomto výskume boli simulované prípady rôznych počtov koncových IoT zariadení, ktorý prijímali dátové pakety. Navrhnuté scenáre boli simulované v simulačnej oblasti bez prekážok o

veľkosti 600 m x 600 m s User Datagram Protocol (UDP) transportným protokolom a veľkosťou paketu 512 bajtov. V rámci každého simulačného scenára, ktorý bol spustený 200 krát sa koncové zariadenia náhodne pohybovali v simulovanej oblasti. Použité parametre s ich hodnotami sú znázornené v Tabuľka 1.

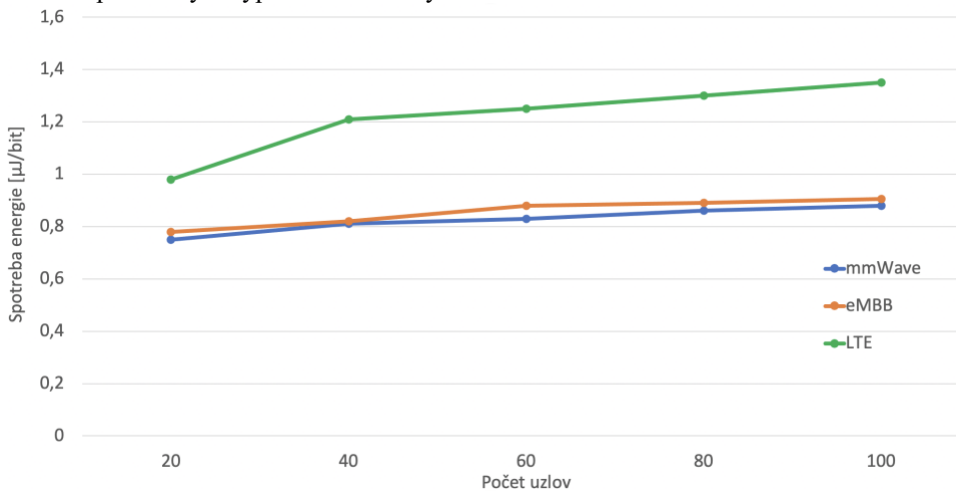
Tabuľka 1 Simulačné parametre

Parameter	Hodnota
Frekvencia	2,160 GHz (LTE), 6 GHz (eMBB), 30 GHz (mmWave)
Šírka pásma	20 MHz
Počet IoT uzlov	20, 40, 60, 80, 100
Mobilita koncových užívateľov	Náhodná
Rýchlosť koncových užívateľov	1-5 km/h
Rozmer simulovanej oblasti	600 x 600 m
Simulačný čas	100 s
Počet simulácií	200
Transportný protokol	UDP
Veľkosť paketu	512 bajtov

### B. Dosiahnuté výsledky

V tejto časti sú opísané dosiahnuté výsledky simulácií, ktoré sú založené na energetickej účinnosti a miere spotrebovanej energie v prostredí simulátora NS-3. Výsledky boli porovnané pre LTE, eMBB a eMBB s mmWave.

Na Obr. 3 nižšie, sú znázornené priemerné hodnoty spotrebovanej energie pre všetky typy simulovaných scenárov. Graf spotrebovanej energie v závislosti od počtu koncových užívateľov ukazuje, že zo zvyšujúcim sa počtom koncových užívateľov sa spotreba energie na bit zväčšuje. Z pohľadu spotrebovanej energie sú služby 5G sietí eMBB a mmWave viac úsporné ako staršia služba LTE pri všetkých typoch simulovaných scenárov.



Obr. 3 Priemerné hodnoty spotreby energie

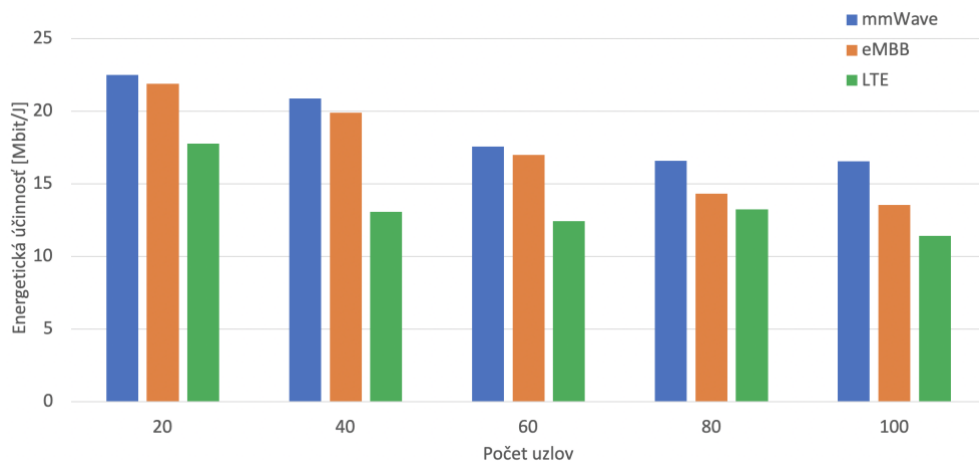
Môžeme vidieť, že priemerné hodnoty spotrebovanej energie LTE služby sa pohybujú od 0,98 μJ/bit do 1,38 μJ/bit v závislosti od počtu koncových užívateľov. Pri porovnaní so službami eMBB a eMBB s mmWave je to výrazne viac. Služba eMBB dosahovala hodnoty od 0,76 μJ/bit do 0,905 μJ/bit. Služba mmWave dosiahla podobné hodnoty, od 0,72 μJ/bit do 0,88 μJ/bit. Všetky priemerné hodnoty spotrebovanej energie sú znázornené v Tabuľka 2 nižšie.

Tabuľka 2 Namerané hodnoty spotreby energie pre všetky scenáre

	20 Koncových užívateľov	40 Koncových užívateľov	60 Koncových užívateľov	80 Koncových užívateľov	100 Koncových užívateľov
<b>LTE</b>	0,98	1,21	1,25	1,3	1,35
<b>eMBB</b>	0,78	0,82	0,88	0,89	0,905
<b>mmWave</b>	0,75	0,81	0,829	0,86	0,88

Ak porovnáme mieru spotrebovanej energie služieb eMBB a mmWave vzhľadom k službe LTE zistíme, že eMBB spotrebovala o 29,81% a mmWave o 32,22% menej energie ako LTE.

Z pohľadu energetickej účinnosti siete boli dosiahnuté výsledky veľmi podobné. Na Obr. 4 má staršia služba LTE nižšiu energetickú účinnosť ako služby eMBB a mmWave. Priemerné hodnoty boli od 17,75 Mbit/J do 11,42 Mbit/J. Služba eMBB dosahovala hodnoty od 33,45 Mbit/J do 13,53 Mbit/J. Služba eMBB s mmWave dosiahla priemerné hodnoty o 1 Mbit/J vyššie ako služba eMBB v rámci všetkých simulovaných scenárov.



Obr. 4 Priemerné hodnoty energetickej účinnosti

Všetky dosiahnuté výsledky energetickej účinnosti sú zobrazené v Tabuľka 3 nižšie.

Tabuľka 3 Namerané hodnoty energetickej účinnosti pre všetky scenáre

	20 Koncových užívateľov	40 Koncových užívateľov	60 Koncových užívateľov	80 Koncových užívateľov	100 Koncových užívateľov
<b>LTE</b>	17,75	13,06	12,43	13,22	11,42
<b>eMBB</b>	21,89	19,88	16,98	14,33	13,53
<b>mmWave</b>	22,49	20,87	17,56	16,58	16,54

#### IV. ZÁVER

Táto práca sa zaoberá komunikáciou pomocou milimetrových vln v 5G sieťach a porovnaním tejto služby so službami enhanced Mobile BroadBand a LTE v oblasti energetickej účinnosti a spotrebovanej energie IoT uzlov.

V tejto štúdií sme sa zamerali na mmWave kvôli tomu, že v budúcnosti bude mať mmWave najväčší dopad na prenosové siete z hľadiska šírky pásma, prenosovej rýchlosti a energetickej účinnosti v porovnaní s eMBB alebo staršou službou LTE.

Na základe simulácií 5G a 4G sietí pre službu UDP boli zanalyzované dosiahnuté výsledky energetickej účinnosti a spotreby energie IoT uzlov v sieti. Jednotlivé scenáre boli simulované pomocou simulátora NS-3 do ktorého bol implementovaný mmWave a energetický model. Na základe dosiahnutých výsledkov sa dá povedať, že priemerné hodnoty energetickej účinnosti a spotrebovanej energie pre 5G siete v porovnaní so 4G sieťami dosahujú lepšie hodnoty. To znamená, že 5G siete sa ukazujú ako efektívnejšie nie len z hľadiska prenosovej rýchlosti, kapacity kanála alebo efektivity prenosu dát, ale aj z pohľadu energetickej účinnosti.

Dosiahnuté výsledky budú tvoriť teoretický základ v budúcej práci, v ktorej sa zameriame na využitie algoritmov strojového učenia s cieľom zvýšiť robustnosť 5G siete, ktorá bude

experimentálne simulovaná v reálnych podmienkach pomocou inteligentných zariadení.

#### POĎAKOVANIE

Výskum prezentovaný v tomto príspevku bol podporený Ministerstvom školstva, vedy, výskumu a športu Slovenskej republiky v rámci výskumného projektu VEGA 1/0753/20, KEGA 009TUKE-4/2019, a Slovenskou agentúrou pre výskum a vývoj s projektom APVV-17- 0208.

#### LITERATÚRA

- [1] N. Baldo, M. Miozzo, M. Requena-Esteso and J. Nin-Guerrero, "An open-source product-oriented LTE network simulator based on ns-3", Proc. 14th ACM Int. Conf. Model. Anal. Simulat. Wireless Mobile Syst., pp. 293-298, 2011, [online] Available: <http://doi.acm.org/10.1145/2068897.2068948>, in press.
- [2] TE-EPC Network Simulator (LENA), Feb. 2012, [online] Available: [http://iptechwiki.cttc.es/LTE-EPC\\_Network\\_Simulator\\_\(LENA\)](http://iptechwiki.cttc.es/LTE-EPC_Network_Simulator_(LENA)), unpublished.
- [3] H. Tazaki et al., "Direct code execution: Revisiting library OS architecture for reproducible network experiments", Proc. 9th ACM Conf. Emerg. Netw. Exp. Technol. (CoNEXT), pp. 217-228, 2013, [online] Available: <http://doi.acm.org/10.1145/2535372.2535374>, in press.
- [4] M. Rebato, M. Zorzi, "Simulation analysis of algorithms for interference management in 5G cellular networks using spatial spectrum sharing", Universita Degli Studi di Padova, pp. 17-33, 2015.
- [5] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, S. Rangan, "5G mmWave Module for ns-3 Network Simulator", NYU Polytechnic School of Engineering, New York, June 2015, in press.
- [6] Edoardo B., "Design and performance evaluation of mm-wave vehicular networks", Politecnico di Torino, pp. 38-44, March 2019.
- [7] A. A. Oliveira, D. Batista, R. Hirata Jr, "Exploring the ns-3 mmWave module", April 2019, [online] Available: <http://vision.ime.usp.br/~arturao/mmwave/mmWaveReport.pdf?fbclid=IwAR3ubQX9V7hrcC546eOvavTNXXBuU9TsMdPjnLy7BKRbXNb54spizqMkLxI>.
- [8] The LENA ns-3 LTE Module Documentation, Jan. 2014, [online] Available: [http://iptechwiki.cttc.es/LTE-EPC\\_Network\\_Simulator\\_\(LENA\)](http://iptechwiki.cttc.es/LTE-EPC_Network_Simulator_(LENA)), unpublished.
- [9] M. M. A. Rojas, I. Demirkol, "Analysis of TCP performance for LTE- 5G Millimeter Wave Dual Connectivity", UPC, Barcelona, Oct. 2019, [online] Available: <https://upcommons.upc.edu/bitstream/handle/2117/174568/Merly+Ascion+Rojas's+Thesis.pdf;jsessionid=E00D80A41D0CCBE0CECF13D40ECFAF4?sequence=4&fbclid=IwAR2bUWfvx8QraKc81OBomGfIf6xGldESnqnVu4vlaVFTmp6Xdkxcp4GEU>.
- [10] "The ns-3 energy framework." [Online]. Available: <http://www.nsnam.org/docs/release/3.20/models/html/energy.html>
- [11] C. Tapparello, H. Ayatollahi, W. Heinzelman "Energy Harvesting Framework for Network Simulator 3 (ns-3)", University of Rochester, Rochester, New York, November 2014.
- [12] "Energy model" [Online]. Available: [https://www.nsnam.org/wiki/Energy\\_model](https://www.nsnam.org/wiki/Energy_model)

# Concept of Concentration Metric for Convolutional Neural Networks

<sup>1</sup>Ivan ČÍK, <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>ivan.cik@tuke.sk, <sup>2</sup>marian.mach@tuke.sk

**Abstract** – Significant changes in hardware development have enabled the scientific community to develop deep learning algorithms. Although these algorithms achieve admirable performance, they are too complex for a developer to understand in detail the internal processes in such an algorithm. This challenge has allowed the rebirth of explainable artificial intelligence. In this article, we present results from experiments performed with Grad-CAM and LRP methods on CNN models trained on MNIST, MNIST-Fashion, and Chicago Face datasets. The aim of the article is to introduce a concept for the model concentration metric, which can help evaluate the quality of the model and lead to improved accuracy of the classifier in the future.

**Keywords** – Convolutional neural networks, Explainable artificial intelligence, Grad-CAM, LRP

## I. INTRODUCTION

The increase in computational power has allowed the emergence of a sub-area of artificial intelligence called deep learning and we also achieved super performance in many tasks that were previously thought to be computationally unattainable [1]. In some areas of daily life, such as healthcare [2], finance [3], and social media [4], ML models are increasingly being used. The outputs of these models can be part of decision-making systems, where the individual actions can affect the human user life, accountability concerns are becoming more and more relevant. Questions for system designers, such as: Does the system work as intended?, Do the decisions being made seem reasonable? or Are we consistent with equality regulation and legislation? are critical, while an entity that may be affected by such an algorithm may be more concerned with topics such as Am I being treated fairly? or What should I do differently next time to get an advantageous outcome [5]? These problems are not unique to computerized decision-making systems, but they have become much more critical with the growth of ML-based systems [6], [7], [8], [9], [10], [11]. The key solution would be to know why the algorithm output is as trustworthy as obtaining this information would not reduce the performance of the algorithm. Due to these problems, **explainable artificial intelligence (XAI)** has become an area of interest in the research community. One of the most striking factors highlighting the importance of XAI research is how many different groups of people (e.g. lawyers, regulators, ML specialists, philosophers, or futurists) consider this area necessary. However, in order to be able to have a conversation with such different groups, it is first necessary to find out whether they understand the basic concept of AI and what the concept of XAI really represents [12]. Explanations should make the AI algorithm expressive to improve human understanding, confidence in decision making, and promote impartial and just decisions. Thus, in order to maintain transparency, trust, and fairness in the ML decision-making process, an explanation or an interpretable solution is required for ML systems. In Section II we briefly describe the theory of XAI. In Section III the concept of XAI visualization methods is explained. Later, we describe Gradient-weighted Class Activation Mapping (Grad-CAM) and Layer-wise Relevance Propagation (LRP) methods in more detail. In Section IV, we describe experiments that were performed on trained deep neural models on the MNIST, MNIST-fashion, and Chicago Face Datasets. The explanation of the models was performed using LRP and Grad-CAM methods. Subsequently, the individual results are interpreted on the basis of which the concept of concentration metrics is presented in Section V.

## II. EXPLAINABLE ARTIFICIAL INTELLIGENCE

Nowadays, in the field of XAI, scientists are trying to make their algorithms more understandable. But let's stop at the first question, what explanation really is? The term **explanation** can be understood differently according to the field (law, cognitive science, philosophy, and social sciences), where it has a relatively well-defined technical meaning and a plethora of research on types of explanations, their purposes, and their social and cognitive function [5].

One of the critical problems in the application of deep neural networks in real life is their black box nature, which raises ethical and legal issues for users and a lack of trust. XAI is a sub-area of artificial intelligence provides tools, techniques, and algorithms that can provide high-quality interpretable, intuitive and human-understandable explanations of black box models. Based on scientific articles in the XAI area we have defined the following challenges:

- 1) **General Theory of Model Explanation:** Systematic general theory for explaining opaque deep neural networks is a current challenge in the scientific community [13], [14]. There are several issues to be raised in this question: *“What needs to be explained? When is it right to explain? Is it different to explain according to the subject who explains? Does the explanation change over time (never before seen occasions)? What explanation is sufficient?”*. Systematic theory should be useful for model explanation research studies, and once established, some current challenging XAI issues may be able to be solved properly by naturally, and some novel scientific directions may be proposed based on the systematic theories.
- 2) **Effective Explanation:** Depending on the speed of response required to provide an explanation, time can become a critical metric [15], [16] and can be used as one of the factors influencing the quality of the explanation.
- 3) **Methods for Debugging CNN Models:** The explanations are also useful for the scientific community, for example, an explanation can help developers debug a model, which is important for improving it. Based on our research, we can say that the number of methods for debugging convolutional neural networks (CNNs) is limited. This direction is one of the most important that can bring us closer to achieving an XAI. The development of high-quality debugging tools for CNN would help developers in the field of AI so that they would be able to solve model errors and understand model behavior.
- 4) **Building User Trust:** Using adversarial examples [17], [18], [19], perturbation methods and poisoned training datasets [20] can detune model decisions. DNN's inability to face such attacks raises questions of trustworthiness for users. The introduction of metrics for model consistency could help in this issue. AI algorithms are increasingly being implemented in the real world. As the trend increases, so does the demand for building user trust in such solutions, even more so in areas where undesired model decisions can affect a person's life. Addressing this issue opens up a new thread of research that is just beginning to develop [21], [22].

## III. VISUALIZATION METHODS FOR EXPLAINING DNNs

Output from XAI visualization method (see Fig. 1) is the identification of areas (pixels at the input, features or areas) from the input that significantly affect the outputs of the model (e.g. saliency maps).

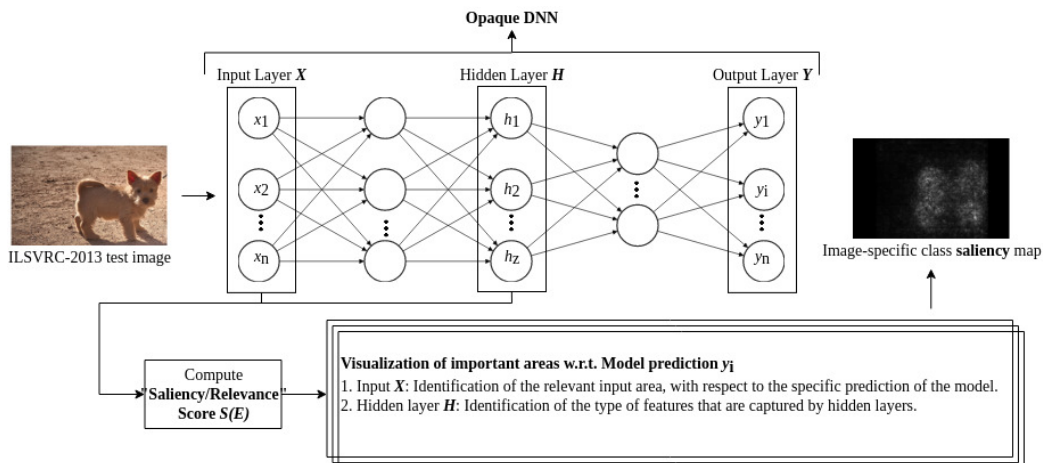


Fig. 1 Extended concept of visualization methods [23].

The  $S(E)$  function is used to calculate the saliency score (see Fig. 1), this function differs according to the visualization method used. The importance for an individual element  $E$  can be calculated for input  $X$  or hidden state  $H$ . An example in the picture is a test picture of a dog from ISLVR-2013, where the output is an image-specific class saliency map. The map was extracted once using back-propagation through the ConvNet classification [24].

#### A. Grad-CAM

The gradient-based interpretability method uses the reverse transmission of information in the neural network to understand the influence of neurons and the correlation between input  $x$  and output  $y$ . Grad-CAM [25] is a class-discriminative attribution technique used to distinguish the neural activity of CNN neurons. Grad-CAM uses the gradients of any class-specific image that flows into the final convolutional layer to generate a map, highlighting the important areas in the image which contribute to the particular model output.

With Grad-CAM it is possible to get the importance map without re-training or changing the base model. To obtain the class-discriminative localization map Grad-CAM  $L_{Grad-CAM}^C \in R^{u \times v}$  of width  $u$  and height  $v$  for any class  $c$ , we first compute the gradient of  $y_c$  (score for any class  $c$ ) with respect to feature maps  $A^k$  of a convolutional layer, i.e.  $\frac{\partial y_c}{\partial A^k}$ . To obtain the neuron importance weights  $\alpha_k^c$ :

$$[\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y_c}{\partial A_{ij}^k}] \quad (1)$$

the gradients that are flowing back are global-average-pooled over the width and height dimensions (indexed by  $i$  and  $j$ ). Grad-CAM heat-map is a weighted combination of feature maps, but followed by a ReLU.

$$[L_{Grad-CAM}^C = ReLU(\sum_k \alpha_k^c A^k)] \quad (2)$$

which will remove the pixels with values less than 0.

#### B. Layer-Wise Relevance Propagation:

In this Section, we describe the layer-wise relevance propagation (LRP) technique introduced in [26]. LRP explains deep neural network prediction and is one of the relevance propagation methods so that the prediction propagates backward without the use of gradients. The algorithm starts from output layer  $L$  of the deep neural network, and then moves in the opposite direction in the graph, gradually redistributing the prediction score until the input is reached. Each neuron receives a share in the output of the network and further redistributes it to its predecessors in the same amount until the input variables are reached [26], [27].

The algorithm does not use transitions and is therefore applicable to general architectures (including networks with discontinuous units). The LRP globally explains the classification decision and the heatmap score has a clear interpretation as evidence for or against the category [27].

LRP is a method of identifying important pixels by inverting a neural network. The reverse pass is a conservative redistribution process, in which the neurons that contribute the most to the previous layer are the most important.

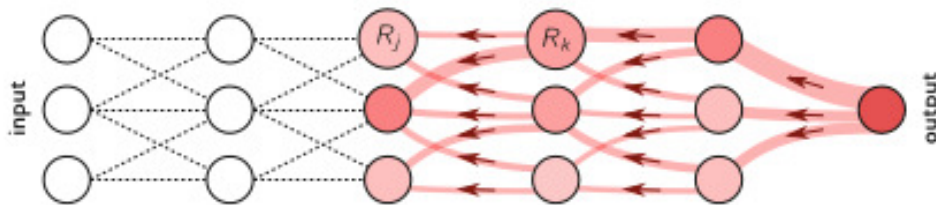


Fig. 2 Diagram of the LRP process. Each neuron is redistributed to lower layers as much as it receives from higher layers [28].

Mathematically, it redistributes the prediction  $f(x)$  backward, using redistribution rules, until it assigns a relevance score  $R_i$  to each variable. An important feature of the redistribution process is called global conservation property and is given as:

$$\sum_i R_i = \dots = \sum_j R_j = \sum_k R_k = \dots = f(x) \quad (3)$$

The described property says that in each step of the redistribution process the overall relevance is maintained. The relevance score  $R_i$  of each input variable determines the extent to which this variable has contributed to the prediction [26].

The global conservation principle is that it preserves the propagated quantity between the neurons of two adjacent layers. When we denote  $R_i^{(l)}$  as the significance associated with the  $i$ -th neuron of layer  $l$  and  $R_j^{(l+1)}$  as the significance associated with the  $j$ -th neuron in the next layer, the conservation principle requires that

$$\sum_i R_i^{(l)} = \sum_j R_j^{(l+1)} \quad (4)$$

where the sums pass through all the neurons of the layers of the given neural model. Applying this rule repeatedly for all layers will ensure that the heatmap resulting from the LRP complies  $\sum_p h_p = f(x)$ , where  $h_p = R_p^{(1)}$  [26].

The LRP method has various modifications such as LRP<sub>2</sub>, LRP<sub>α1</sub>, GB, PA, LRP<sub>CMP</sub>, and others. The methods and rules that apply to them are described in [29].  $\epsilon$ -LRP is used mostly (also in our experiments, see Section IV), which proceeds layer by layer, redistributing prediction scores until it reaches the input layer. It starts at the last layer  $L$ , and the relevance of the target neuron  $c$  is equal to the output of this neuron. This relevance is further redistributed. The relevance of the other neurons in the output layer is zero.

$$r_i^L = \begin{cases} S_i(x) & \text{if unit } i \text{ is the target unit of interest} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

A recursive rule that redistributes the relevance of one layer to the next layer is the  $\epsilon$ -rule

$$r_i^{(l)} = \sum_j \frac{z_{ji}}{\sum_{i'} (z_{ji'} + b_j) + \epsilon * \text{sign}(\sum_{i'} (z_{ji'} + b_j))} r_j^{(l+1)} \quad (6)$$

where we define  $z_{ji}$  to be the weighted activation of a neuron  $i$  onto neuron  $j$  in the next layer and  $b_j$  the additive bias of unit  $j$ . A small amount of  $\epsilon$  is added to the denominator to avoid numerical instability [30].

#### IV. EXPERIMENTS

Experiments were written in Python using PyTorch which is an open source machine learning library [31] that was built to allow for quick development of machine learning solutions, providing many abstract methods and objects to specialize into the desired parts.

##### A. LRP experiments with MNIST and MNIST-fashion dataset

In the first experiment, we used  $\epsilon$ -LRP method. In our experiments, we used the well-known MNIST dataset [32], which consists of handwritten numbers from 0 to 9, so it contains 10 classes. The dataset contains 60,000 training and 10,000 test images, each with the resolution with 28 x 28 x 1 pixels. The neural network model had 2 hidden layers, each with 256 neurons. On the testing dataset, we achieved an accuracy of 96.75%. We explained the model using the  $\epsilon$ -LRP method.

[28].

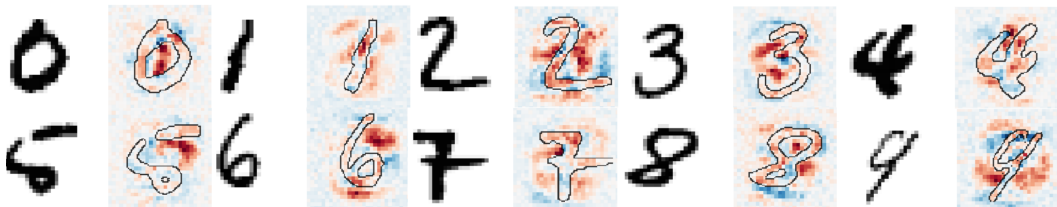


Fig. 3 Original Images from MNIST dataset paired with their heatmap output from  $\epsilon$ -LRP method [23].

In the picture (Fig. 3), we can see the original picture numbers paired with their heatmap map where the individual pixels are marked. Pixels are red if their value is positive, or blue if their value is negative. The intensity of the color depends on the distance of value from zero. Dark red pixels had the greatest influence on the neural network decision when choosing a class.

In the second experiment, we used MNIST-fashion dataset, which consists of 10 classes (t-shirt, trouser, pullover, dress, coat, sandal, shirt, sneaker, bag, and ankle boot). This dataset was more



complex and harder to learn as we achieve an accuracy of 89.12%. First, we took all the test images that belong to the "Coat" class. For each image in this class, we used the LRP method and added the individual values at the same positions, and then divided them by the number of images (see Fig. 4).

The problem of misclassification comes when we have similar heatmaps for two classes, e.g. dog breed. If our accuracy is not high enough, we would look at classes that have very similar cumulative heatmaps, and in the training pictures, we would try to cover the areas on which they learn. In this way, we could get our model to consider other areas than identifying the class. In our opinion, the LRP method will not satisfy the classical user's explanations, but it can help the developer to understand the behavior of the model.

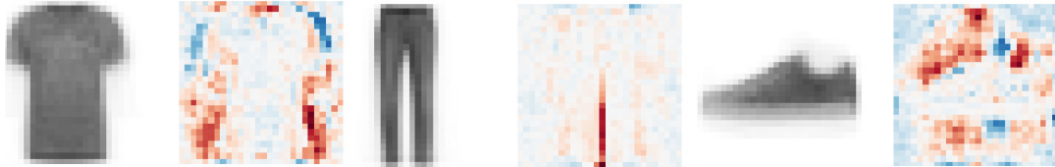


Fig. 4 Cumulative heatmap for classes t-shirt, trouser and sneaker.

### B. LRP and Grad-CAM experiments on Chicago Face Dataset

A group of scientists in the field of psychology and the study of emotions has compiled the Chicago Face Database [33]. It is a freely available database consisting of high-resolution, standardized photos of 158 subjects of various ethnicity aged 17 to 65 years, expressing different emotions. Each face in the database had one expression among five: **neutral**, **angry**, **fearful**, **happy with mouth closed**, **happy with mouth opened**. Each image has a resolution of 2444 (wide) by 1718 (high) pixels, making really small details recognizable. The high-resolution photographs in the Chicago Face Dataset would cause our network to train too long and be sensitive to small details, so to avoid these problems, all the images were resized to 128 by 128 pixels and normalized using the mean and standard deviation of the dataset. That way the data will be processed faster and the training will be easier, while still having enough salient features on the images to be relevant.

The neural network that was used in experiments in itself was made of 5 layers. The first 3 layers are intertwined convolution layers and batch normalization layers. Finally, the last 2 layers are fully dense layers used for the classification of features learned by the convolutions. For the training of the network, a learning rate of 0.001 was chosen after some trial and error. The loss was calculated using the negative log-likelihood function of PyTorch. The final accuracy acquired on the validation dataset were of around 65% across 5 different runs. To add a measure of understanding to the artificial neural network, we decided to add a visualization of the areas that hold high importance in the decision process of the neural network. To accomplish that, we relied on the Grad-CAM and LRP (see Fig. 5).

[28].

### C. Results

Our analysis showed that it did so by focusing on parts of the images that are not so important. This mainly shows that our networks try to identify patterns that are not important. Another hypothesis and more probable is that the network identified and learned a bias present in the data that the annotator nor we humans did not identify. Such bias generally harms the performance of Deep Neural Networks when used in a real-world application. Therefore choosing the right dataset and transformation is useful for DNN networks to properly learn the foreground pattern. In future works, we will investigate models trained on wild datasets such as AffectNet.

The GRAD-CAM and LRP methods did not add much insight into the process of the artificial neural network either. The result of Grad-CAM and LRP on Chicago Face Dataset (see Fig. 5) of that endeavor is something that we were not expecting: the heat-map displays the hair and shoulders as being the highest importance for the network, with the facial features coming after in intensity. As an output, we would expect the class-discriminative localization map to be located mainly in the facial area, as we trained our network to recognize emotions using facial expressions. Despite the fact that our trained Deep Neural Networks correctly classified those images in Fig. 5 with very high accuracy.

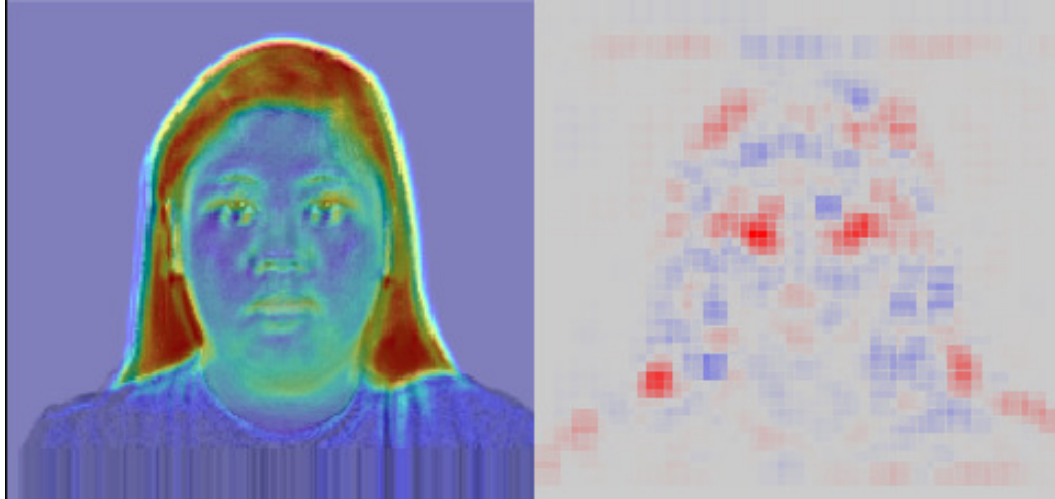


Fig. 5 Correctly classified neutral face expression sample explained with Grad-CAM and LRP.

### V. CONCEPT OF CONCENTRATION METRIC

Introducing a concentration metric and further research in this area could lead to increasing the generalization of CNN so that the classification is not sensitive to random features that do not actually describe the object. By comparing the class activation map (evaluated on different visualization XAI approaches) and segmentation of an object from an image, we can identify whether the model focuses on the correct parts of the image. By using the proper loss function, we should be able to maintain the model in the areas that matter by punishing neurons that should not be activated highly. Introducing a metric for concentrating a model on a given class can help us describe the quality of the model. As shown in Fig. 6 we can suppose an input image  $x_j$ . Using XAI visualization methods to explain a CNN we can observe an output in form of map  $m_j$  which describes the importance of individual pixels (, LRP identify pixels that have had a negative or positive effect on image classification).

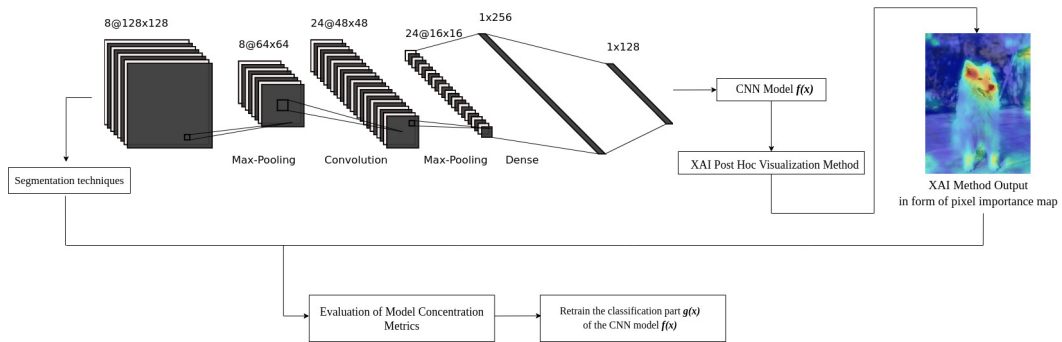


Fig. 6 Concept of concentration metric.

Based on the experiments from the previous Section IV, we describe the way in which we would like to implement a given metric:

In the input image  $x_j$ , we segment the object  $o_j$  of the given class. To calculate the degree of concentration, assume that:

- $O$  is the sum of pixels in the area where the segmented object  $o_j$  from the input image is located.
- $O^+$  is the sum of positively activated pixels in the segmented object area  $o_j$  generated at the output with, for example LRP method.
- $O^-$  is the sum of negative and neutral activated pixels in the segmented object area  $o_j$  generated at the output with, for example LRP method.
- $B$  is the sum of pixels in the image subtracted by  $O$  (= area where the segmented object  $o_j$  is not located). item  $B^+$  is the sum of positively activated pixels in area without the object.
- $B^-$  is the sum of negative and neutral activated pixels in area without the object.

We would like to use these metrics to define an equation for measuring model quality.

## ACKNOWLEDGMENT

This research work was supported by APVV project 015-0730 “Cloud Based Human Robot Interaction” and Feigrant2021.

## REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [2] J. De Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O’Donoghue, D. Visentin *et al.*, “Clinically applicable deep learning for diagnosis and referral in retinal disease,” *Nature medicine*, vol. 24, no. 9, pp. 1342–1350, 2018.
- [3] J. B. Heaton, N. G. Polson, and J. H. Witte, “Deep learning for finance: deep portfolios,” *Applied Stochastic Models in Business and Industry*, vol. 33, no. 1, pp. 3–12, 2017.
- [4] O. Alvarado and A. Waern, “Towards algorithmic experience: Initial efforts for social media contexts,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.
- [5] B. Mittelstadt, C. Russell, and S. Wachter, “Explaining explanations in ai,” in *Proceedings of the conference on fairness, accountability, and transparency*, 2019, pp. 279–288.
- [6] B. Bodó, N. Helberger, K. Irion, F. Zuiderveen Borgesius, J. Moller, B. van de Velde, N. Bol, B. van Es, and C. de Vreese, “Tackling the algorithmic control crisis—the technical, legal, and ethical challenges of research into algorithmic agents,” *Yale JL & Tech.*, vol. 19, p. 133, 2017.
- [7] J. A. Kroll, S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson, and H. Yu, “Accountable algorithms,” *U. Pa. L. Rev.*, vol. 165, p. 633, 2016.
- [8] H. Nissenbaum, “Accountability in a computerized society,” *Science and engineering ethics*, vol. 2, no. 1, pp. 25–42, 1996.
- [9] S. C. Olhede and P. J. Wolfe, “The growing ubiquity of algorithms in society: implications, impacts and innovations,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2128, p. 20170364, 2018.
- [10] F. Pasquale, *The black box society*. Harvard University Press, 2015.
- [11] A. D. Selbst and S. Barocas, “The intuitive appeal of explainable machines,” *Fordham L. Rev.*, vol. 87, p. 1085, 2018.
- [12] Z. C. Lipton, “The mythos of model interpretability,” *Queue*, vol. 16, no. 3, pp. 31–57, 2018.
- [13] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins *et al.*, “Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, 2019.
- [14] J. Díez, K. Khalifa, and B. Leuridan, “General theories of explanation: Buyer beware,” *Synthese*, vol. 190, no. 3, pp. 379–396, 2013.
- [15] M. A. Ahmad, C. Eckert, and A. Teredesai, “Interpretable machine learning in healthcare,” in *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 2018, pp. 559–560.
- [16] S. Nemati, A. Holder, F. Razmi, M. D. Stanley, G. D. Clifford, and T. G. Buchman, “An interpretable machine learning model for accurate prediction of sepsis in the ICU,” *Critical care medicine*, vol. 46, no. 4, pp. 547–553, 2018.
- [17] X. Yuan, P. He, Q. Zhu, and X. Li, “Adversarial examples: Attacks and defenses for deep learning,” *IEEE transactions on neural networks and learning systems*, 2019.
- [18] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” *arXiv preprint arXiv:1412.6572*, 2014.
- [19] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, “Towards deep learning models resistant to adversarial attacks,” *arXiv preprint arXiv:1706.06083*, 2017.
- [20] N. Müller, D. Kowatsch, and K. Böttlinger, “Data poisoning attacks on regression learning and corresponding defenses,” in *2020 IEEE 25th Pacific Rim International Symposium on Dependable Computing (PRDC)*. IEEE, 2020, pp. 80–89.
- [21] H. Jiang, B. Kim, M. Y. Guan, and M. R. Gupta, “To trust or not to trust a classifier,” in *NeurIPS*, 2018, pp. 5546–5557.
- [22] J. Heo, H. B. Lee, S. Kim, J. Lee, K. J. Kim, E. Yang, and S. J. Hwang, “Uncertainty-aware attention for reliable interpretation and prediction,” in *Advances in Neural Information Processing Systems*, 2018, pp. 909–918.
- [23] N. Xie, G. Ras, M. van Gerven, and D. Doran, “Explainable deep learning: A field guide for the uninitiated,” *arXiv preprint arXiv:2004.14545*, 2020.
- [24] K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep inside convolutional networks: Visualising image classification models and saliency maps,” *arXiv preprint arXiv:1312.6034*, 2013.
- [25] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 618–626.
- [26] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek, “On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation,” *PLoS one*, vol. 10, no. 7, p. e0130140, 2015.
- [27] W. Samek, A. Binder, G. Montavon, S. Lapuschkin, and K.-R. Müller, “Evaluating the visualization of what a deep neural network has learned,” *IEEE transactions on neural networks and learning systems*, vol. 28, no. 11, pp. 2660–2673, 2016.
- [28] G. Montavon, A. Binder, S. Lapuschkin, W. Samek, and K.-R. Müller, “Layer-wise relevance propagation: an overview,” in *Explainable AI: interpreting, explaining and visualizing deep learning*. Springer, 2019, pp. 193–209.
- [29] M. Kohlbrenner, A. Bauer, S. Nakajima, A. Binder, W. Samek, and S. Lapuschkin, “Towards best practice in explaining neural network decisions with lrp,” in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7.
- [30] M. Ancona, E. Ceolini, C. Öztireli, and M. Gross, “Towards better understanding of gradient-based attribution methods for deep neural networks,” *arXiv preprint arXiv:1711.06104*, 2017.
- [31] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, “Pytorch: An imperative style, high-performance deep learning library,” in *Advances in neural information processing systems*, 2019, pp. 8026–8037.
- [32] Y. LeCun, “The mnist database of handwritten digits,” <http://yann.lecun.com/exdb/mnist/>, 1998.
- [33] D. S. Ma, J. Correll, and B. Wittenbrink, “The chicago face database: A free stimulus set of faces and norming data,” *Behavior research methods*, vol. 47, no. 4, pp. 1122–1135, 2015.

# Virtuálne prostredie pre podporu vzdelávania hendikepovaných osôb

<sup>1</sup>Branislav SOBOTA, <sup>2</sup>Štefan KOREČKO, <sup>3</sup>Matej MASRNA

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

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

**Abstrakt** — Tento článok opisuje aplikáciu určenú na vzdelávanie žiakov so špeciálnymi výchovno-vzdelávacími potrebami. Technológie virtuálnej reality (VR) a najmä webovej VR sú postupne čoraz viac využívané v rôznych oblastiach našej spoločnosti, aj v súvislosti s pandemickou situáciou a využívaním dištančného vzdelávania. V rámci článku je opísané aké požiadavky je potrebné pri tvorbe vzdelávacej aplikácie pre žiakov so špeciálnymi potrebami splniť a akým spôsobom to ovplyvní používateľské rozhranie aplikácie. Článok sa venuje aj vytvoreným virtuálnym prostrediam a ich možnému využitiu nielen vo vyučovacom procese.

**Kľúčové slová** — webová virtuálna realita, hendikepované osoby, dištančné vzdelávanie

## I. ÚVOD

Virtuálna realita [1] je čoraz častejšie využívaná v rôznych oblastiach našej spoločnosti. Ide o oblasti ako veda, výskum, zábava, alebo architektúra. Jednou z oblastí, v ktorej vidíme potenciál využitia virtuálnej reality na Slovensku je podľa nášho názoru školstvo a vzdelávanie. Vzdelanie je považované za jednu z najdôležitejších potrieb. Nárok na vzdelanie má každé dieťa. Pristupuje však naša spoločnosť k hendikepovaným deťom dostatočne kvalitne? Vo väčšine prípadov sú hendikepované deti súčasťou špeciálneho vzdelávania. Špeciálne vzdelávanie na Slovensku sa však stretáva s viacerými problémami. Jedným z problémov sú učebné materiály pre žiakov. Vzhľadom na špecifiká prostredia nie je ľahké vytvoriť učebné materiály z dôvodu veľkej miery individuálnych potrieb. Problém však môže byť aj s dištančným vzdelávaním žiakov, čo koniec-koncov preukázala aj súčasná pandemická situácia. Aplikácie využívané na bežných školách môžu byť pre žiakov s hendikepom zbytočne zložité a nepochopiteľné. Tento problém by teda mohol pomôcť vyriešiť systém, ktorý by bol pre žiakov jednoduchý, k dispozícii aj v domácich podmienkach a bol by prispôsobený ich potrebám a schopnostiam.

Pri špeciálnom vzdelávaní sa stretávame so žiakmi, ktorí majú ťažkosti s učením. Ako je uvedené v [2] platí, že pojem ťažkosti s učením nesúvisí so sociálnymi a ekonomickými faktormi, ale biologickým poškodením. Biologické poškodenia u daných detí majú vplyv na vývoj a spôsobujú obmedzenia v rôznych oblastiach. Okrem toho však môžu spôsobovať znížené schopnosti žiaka pri vzdelávaní. Medzi ťažkosti s učením je možné zaradiť pohybové postihnutie, zmyslové poruchy, kognitívne poruchy ako mentálna retardácia, rôzne druhy autizmu, nedostatok pozornosti alebo poruchy správania. Takéto deti môžu mať problémy aj v rámci sociálnej interakcie alebo napríklad predstavivosti. Navyše je potrebné prihliadať na technologické možnosti a obmedzenia pri nasadzovaní moderných technológií akou je aj virtuálna realita a s ňou súvisiace technológie.

Z hľadiska postihnutia a teda následných potrieb takýchto ľudí v zmysle ich používateľského rozhrania sa evidujú nasledujúce skupiny: *telesné postihnutie* (poruchy jemnej a hrubej motoriky, znížené mobilné schopnosti, atď.), *zmyslové postihnutie* (znížené zrakové alebo sluchové schopnosti), *mentálne postihnutie* (napr. Downov syndróm), *rečové postihnutie* (znížené rečové schopnosti), *špecifické poruchy správania* (ľahké mozgové dysfunkcie), *špecifické poruchy učenia* (poruchy pamäti, pozornosti, problémy s čítaním, písaním, počítaním, atď.) a iné postihnutie (rôzne druhy závislosti, senilita, a iné).

Informačné technológie môžu pomôcť zdravotne postihnutým, najmä v týchto oblastiach [3]:

- Uľahčenie prístupu k informáciám,
- Pomoc pri odstraňovaní defektov,
- Uľahčenie vykonávania bežných činností,
- Umožnenie komunikácie s intaktnými ľuďmi,
- Príprava študijných materiálov a učebných pomôcok pre zdravotne postihnutých.

Ludia trpiaci určitým hendikepom sú schopní osvojiť si rôzne zručnosti a návyky, ak sa im venuje dostatok trpezlivosti a času. S použitím virtuálno-reálnych technológií, ktoré interaktívne a vizuálne kvalitnejšie spracujú učivo v rámci výchovno-vzdelávacieho procesu, sa predpokladá rýchlejšie, dokonalejšie a ľahšie pochopenie učiva. Dané učivo je potrebné prispôbiť konkrétnemu hendikepu. Úplne novou oblasťou je implementácia hlavne kolaboratívnej VR a pridružených technológií [4], ktoré najmä v čase pandémie, môžu byť nápomocné pri dištančnej výuke.

Hlavným pilierom rozhrania pre hendikepované osoby je prístupnosť. Do veľkej miery závisí od typu postihnutia. Preto je analýza orientovaná na typy postihnutia, pre ktoré Spojená škola Pavla Sabadoša internátna v Prešove poskytuje vzdelanie. Škola je špecializovaná pre žiakov s mentálnym postihnutím, viacnásobným postihnutím, autizmom a kombinovaným postihnutím telesným a mentálnym. A využitie technológií VR a aj zmiešanej reality [5] bolo predmetom projektu, ktorý bol riešený s touto školou. Podľa [6] niektoré hendikepované osoby môžu mať problém s použitím myši, kvôli obmedzenej kontrole jemnej motoriky. Preto je odporúčané pre túto skupinu používateľov použiť podporné technológie pre vstup ako napríklad vstup reči alebo dotykový monitor. Aplikácie určené pre takéto osoby by sa mali skladať z dostatočne veľkých častí grafického používateľského prostredia pre bezproblémovú manipuláciu s nimi a jasnému čítaniu popisov. Používatelia s kognitívnymi poruchami [7] môžu mať problémy s navigáciou v rozhraní a pri čítaní dlhších textov. Preto je vhodné prostredie maximálne zjednodušiť. Existujúce riešenia sú v prítomnosti prispôbované predovšetkým typom hendikepu jednotlivých osôb. Z tohto pohľadu použitie virtuálno-reálnych technológií môže byť rôzneho charakteru. V nasledujúcom článku bude opísané jedno z riešení, ako sprístupniť pomocou hlavne webových VR technológií, niektoré koncepty vyplývajúce z výuky uvedených žiakov.

## II. NÁVRH SYSTÉMU

Prvá významná požiadavka sa týka dostupnosti riešenia. Implementované riešenie malo byť dostupné žiakom a pedagógom hlavne v školskom prostredí na vyučovacom procese. Okrem toho je tu aj logická požiadavka vytvoriť systém dostupný aj v domácich podmienkach žiakov. Z toho dôvodu je vhodné implementovať virtuálno-reálne technológie priamo vo webovom prostredí. Na implementáciu je vhodné využiť webový rámec virtuálnej reality A-frame. V rámci A-frame je k dispozícii možnosť tvorby kolaboratívnej virtuálnej reality [4] pomocou Networked A-frame, čo prináša ďalšie zaujímavé možnosti pre tvorbu vzdelávacej aplikácie. Kolaboratívna virtuálna realita je založená na nejakej forme spolupráce. Ako uvádza literatúra [8] v rámci kolaboratívneho virtuálneho prostredia platí, že viacerí používatelia môžu interagovať medzi sebou alebo s okolitými objektami vďaka zdieľaniu virtuálneho prostredia. Každý z používateľov je v rámci prostredia reprezentovaný vlastným avатарom.

### A. Názornosť virtuálnych objektov

Názornosť je vyžadovaná hlavne na napĺňanie potrieb žiakov s mentálnym postihnutím, nakoľko títo žiaci majú problém s predstavivosťou. V prípade mentálnej retardácie sa podľa [9] môžeme stretnúť taktiež s výrazným obmedzením pri prispôbovaní sa na bežné životné podmienky a zníženou schopnosťou učenia sa. V oblasti riešenia VR sa názornosť týka hlavne používaných modelov vo virtuálnych prostrediach. Je potrebné, aby modely boli tvarom a inými vlastnosťami realistické, avšak podmienkou nie je ich vysoká presnosť a detaily. Na vytvorenie virtuálnych avatarov (Fig. 1) bol použitý modelovací nástroj Blender, ktorý je voľne dostupný a ponúka zaujímavé možnosti.

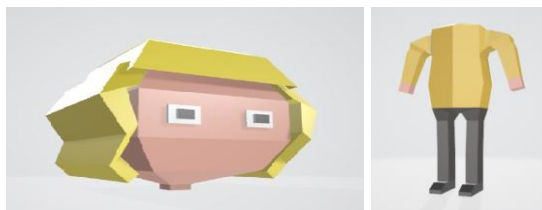


Fig. 1 Hlava virtuálneho avatara učiteľky a telo virtuálneho avatara žiaka

### B. Rýchlosť odozvy rozhrania aplikácie

Ďalšia významná požiadavka návrhu súvisí s rýchlosťou odozvy rozhrania aplikácie z dôvodu sústredenia žiakov. Aplikácia by mala pri používaní pôsobiť čo najplynulejšie. To je možné splniť v prípade využitia webovej virtuálnej reality tým spôsobom, že aplikácia bude pôsobiť ako jednostránková aplikácia. Požiadavke je možné vyhovieť použitím reaktívnych technológií. V opisovanom riešení sme sa pri tvorbe používateľského rozhrania využili knižnica React.js.

### III. ARCHITEKTÚRA CELKOVÉHO RIEŠENIA

Celková architektúra riešenia sa skladá z piatich základných častí (Fig. 2):

- *vstupno-výstupné zariadenie* - osobný počítač bol zvolený vzhľadom na technologické vybavenie slovenských základných škôl a vôbec vybavenie na spojenej škole internátnej.
- *používateľské rozhranie* - využitie React.js, pre grafické úpravy kaskádové štýly v kombinácii s rámcom React Semantic UI
- *webový server* - využitie technológie Node.js
- *databáza predmetov* - využitie technológie Firebase a databázy Firestore
- *repozitár modelov* - repozitár trojrozmerných modelov

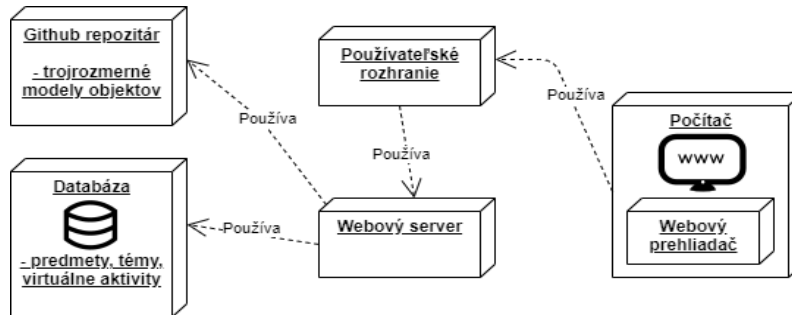


Fig. 2 Celkový návrh architektúry riešenia

### IV. POUŽÍVATEĽSKÉ ROZHRAKIE APLIKÁCIE

Jednou z dôležitých častí nami vytváranej aplikácie určenej na vzdelávanie detí v rámci špeciálneho vzdelávania je grafické používateľské rozhranie. Cieľom bolo vytvoriť používateľské rozhranie, ktoré bude pre žiakov ako aj učiteľov jednoduché na používanie. Pri tvorbe používateľského rozhrania boli zohľadňované aj požiadavky týkajúce sa potrieb a schopností žiakov. Celé grafické rozhranie je založené na dvoch základných princípoch. Medzi tieto princípy patria:

- využívanie obrázkov a ikon,
- využívanie farieb.

#### A. Využívanie obrázkov a ikon

Prvým základným princípom grafického používateľského rozhrania aplikácie je využívanie obrázkov a ikon. Pri tomto princípe sa vychádza z predpokladu, že obrázky a ikony sú výraznejšie oproti textovým prvkom. Každý obrázok v rámci rozhrania má svoje jasné opodstatnenie a účel. Napríklad obrázok s ikonou domčeka (Fig. 3) značí, že po kliknutí na tento obrázok sa používateľ dostane na domovskú obrazovku. Takýmto spôsobom sú v rámci rozhrania používané všetky príslušné obrázky a ikony.



Fig. 3 Rozhranie v rámci virtuálnej miestnosti

#### B. Využívanie farieb

Druhým základným princípom rozhrania aplikácie sú farby. Farby sa využívajú na upútanie pozornosti používateľa/žiaka. Farebné veci priťahujú pozornosť ľudí, avšak musia byť využívané opodstatnene. Výrazné farby preto v rozhraní sa využívajú pri prvkoch, ktoré si vyžadujú nejakú akciu. Aj z toho dôvodu je základné pozadie aplikácie v bielej farbe (tiež aj preto, aby nerozptyľovalo žiakov). Druhým spôsobom akým sú farby v rámci rozhrania využívané, je informačná funkcia. Zelená farba je pri informáciách smerovaných k používateľom využívaná napríklad pri úspešnom zvládnutí úloh zadaného predmetu (Fig. 4).

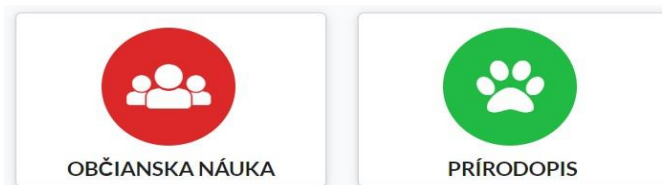


Fig. 4 Farebné karty vyučovacích predmetov

## V. VIRTUÁLNE PROSTREDIA

V rámci aplikácie boli vytvorené virtuálne prostredia, ktoré sú rozdelené na nasledovné kategórie:

- aktivity na vyučovanie,
- testovacie aktivity.

### A. Aktivity na vyučovanie

Prvú kategóriu virtuálnych prostredí predstavujú aktivity využívané na vyučovanie. Tieto aktivity je možné rozdeliť na 2 základné druhy. Prvou možnosťou je skupinové vyučovanie (Fig. 5 vľavo). V rámci skupinového vyučovania je možné pripojenie viacerých používateľov do jednej virtuálnej miestnosti, kde vykonávajú spoločnú aktivitu. Druhou možnosťou aktivít je individuálne vyučovanie (Fig. 5 vpravo), ktoré je určené na individuálnu prácu pedagóga so žiakom (veľmi dôležité v špeciálnej pedagogike). V tomto prípade nie sú virtuálne objekty zdieľané, okrem polohy samotného žiaka. V prípade individuálneho vyučovania je učiteľ obmedzený iba na pohľad zhora, pričom nastavenú kameru dokáže ovládať, rôznym spôsobom približovať a otáčať.



Fig. 5 Ukážka z virtuálneho skupinového vyučovania (vľavo) a z virtuálneho individuálneho vyučovania (vpravo)

### B. Testovacie aktivity

Druhú kategóriu virtuálnych prostredí, predstavujú aktivity využívané napríklad na testovanie. Tieto aktivity je možné využiť na praktické overovanie si vedomostí žiakov, prípadne nadobudnutie nových znalostí. Celkovo základná aplikácia ponúka tri aktivity prezentujúce možnosti využitia technológie virtuálnej reality v prostredí špeciálneho vzdelávania. Aplikácia je dynamická a ponúka taktiež možnosť vytvorenia vlastných nových virtuálnych aktivít.

Implementované názorné aktivity:

- *Nájdí ihličnaté stromy* - úlohou žiakov je nájsť v lese všetky ihličnaté stromy (Fig. 7). V lese sú však aj listnaté stromy. Z toho hľadiska je nevyhnutné, aby žiak dokázal rozoznať ihličnatý druh stromu od toho listnatého. Žiaci označujú objekty pomocou virtuálneho kurzoru. Využitie typu prostredia vo výučbe: rozoznávania objektov zo strany žiaka.

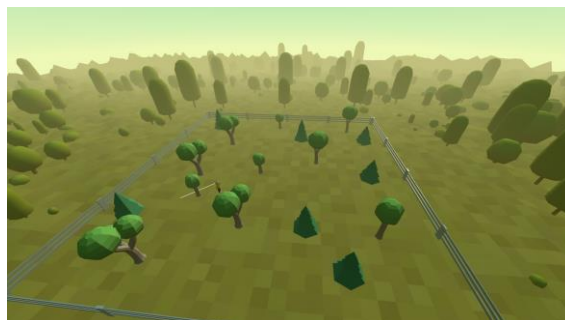


Fig. 7 Aktivita - Nájdí ihličnaté stromy

- *Nájdí cestu von z lesa* - úlohou žiakov je nájsť cestu von z lesa (Fig. 8). Stromy sú však príliš husté na to, aby žiaci dokázali cez ne prejsť. Preto je potrebné hľadať šípky, ktoré ukazujú správny smer. Žiaci sa tým pádom učia chápať význam smeru, rozoznávajú pravej a ľavej strany a podobne. Ak žiaci nasledujú šípky správne, dostanú sa postupne k chalúpke, kde im pomôžu. Využitie typu prostredia vo výučbe: presun vlastnej postavy z bodu A do bodu B, orientácia v priestore

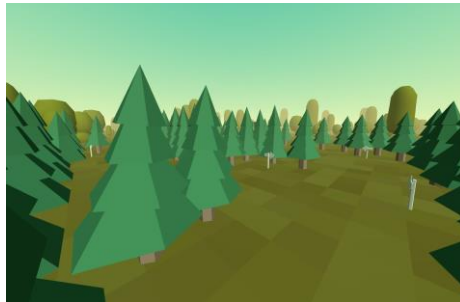


Fig. 8 Aktivita - Nájdí cestu von z lesa

- *Prechod cez cestu* - úlohou žiakov je vo virtuálnom priestore prejsť cez cestu. Po ceste však jazdia autá. Žiak preto musí autá sledovať a rozhodnúť sa, kedy cez cestu prejde. Celé to môže sledovať učiteľ a učiť žiaka, aby sa pozrel na obe strany, prípadne ho vyzval k rozhodnutiu. Využitie typu prostredia vo výučbe: vyhýbanie sa žiakov nebezpečným predmetom alebo živlom, aby správne pochopili ich význam a následky svojej činnosti

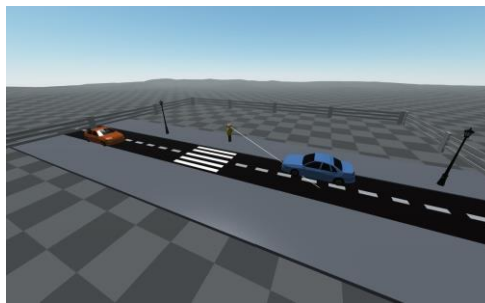


Fig. 9 Aktivita - Prechod cez cestu

Samozrejme popisované aktivity/scenáre sú generické a v zmysle použitia iných virtuálnych objektov/scén je možné dané aktivity vyučovať aj pre iné situácie resp. oblasti.

## VI. TESTOVANIE APLIKÁCIE

V rámci používateľského testovania našej aplikácie z pohľadu žiaka bola využitá technika testovania použiteľnosti. Táto technika je aplikovateľná na našu aplikáciu, pretože účastníci testovania sú deti a školáci. Tým pádom ide o reálnych používateľov typu domény našej aplikácie. Účastníci testovania boli aktuálni, resp. budúci žiaci základnej školy. Vzhľadom na obmedzené podmienky pre testovanie kvôli aktuálne prebiehajúcej epidemiologickej situácii a s tým súvisiacimi opatreniami aj priamo v prostredí špeciálneho vzdelávania sme vybrali ako účastníkov deti, ktoré budú mať, resp. majú skúsenosti z prostredia základnej školy (neboli priamo z prostredia špeciálneho vzdelávania). Vzhľadom na to, že aplikácie je určená primárne pre žiakov 2. stupňa v rámci špeciálneho vzdelávania, berieme nižší vek účastníkov ako zníženie mentálnych schopností, čím chceme simulovať testovanie žiakov s ľahkým stupňom mentálneho postihnutia.

Vzhľadom na zlú pandemickú situáciu v čase vývoja aplikácie sa celkovo testovania aplikácie z pohľadu žiaka zúčastnili iba 2 deti, čo si uvedomujeme, že je málo, avšak vytvoril sa minimálne priestor pre metodiku budúceho testovanie. Vekový rozsah účastníkov bol 6 - 8 rokov. Z toho vyplýva, že priemerný vek účastníkov je 7 rokov. Testovania sa zúčastnil jeden účastník mužského pohlavia a jedna účastníčka ženského pohlavia. Úroveň znalostí účastníkov bola odlišná iba do určitej miery. Prvý účastník testovania je žiakom, ktorý v budúcom školskom roku začne chodiť do 1. triedy na základnej škole. Druhá účastníčka testovania je aktuálne v druhom ročníku na základnej škole. Testovanie, vzhľadom na lockdown, prebiehalo v domácich podmienkach maloletých účastníkov. Pre analýzu výsledkov a zaznamenávanie testovania boli využité nasledovné spôsoby a pomôcky:



- snímanie obrazovky počítača,
- zvukový záznam pomocou diktafónu,
- pero a papier určené na poznámky osoby zodpovednej za testovanie,
- stopky určené na zaznamenávanie času

Pred samotným testovaním boli účastníkom vysvetlené pokyny. Každý z účastníkov vykonával testovanie samostatne iba v prítomnosti zodpovednej osoby. V prípade otázok mali účastníci možnosť sa osoby zodpovednej za testovanie (spravidla rodič) pýtať čo ich zaujíma, prípadne čo im nie je jasné. Po vysvetlení všetkých nezrovnalostí mohlo samotné testovanie začať. Testovanie z pohľadu žiaka bolo rozdelené na štyri časti. Prvá časť bola zameraná na testovanie používateľského rozhrania aplikácie. V druhej časti bolo úlohou účastníkov prejsť 3 aktivity vytvorené výlučne na spoznanie ovládania a pravidiel virtuálnej školy. Tretia časť v sebe zahŕňala virtuálnu aktivitu, pri ktorej mal účastník voľnosť pri používaní. Po uplynutí stanoveného časového limitu bolo úlohou účastníka vymenovať virtuálne objekty a pojmy, s ktorými sa vo virtuálnej miestnosti stretol. Týmto spôsobom sme sa snažili overiť výukovú časť aplikácie. Posledná časť testovania aplikácie z pohľadu žiaka bola zameraná na získanie spätnej väzby zo strany účastníka. Za tým účelom boli účastníkom položené otázky, na ktoré mali odpovedať.

V rámci položených otázok išlo o získanie spätnej väzby na testovanú aplikáciu zo strany detí. Za účelom štandardizovania merania použiteľnosti aplikácie sme využili štandardnú škálu pre hodnotenie v rámci techniky SUS (System usability scale). V rámci tohto systému bolo celkovo účastníkom položených 10 štandardných otázok. Pri každej otázke bolo úlohou účastníka vybrať jednu z piatich možností na bodovej škále 1 – 5 (formou etikonov). Vzhľadom na to, že účastníkmi nášho testovania boli deti, bolo potrebné upraviť štandardné otázky v rámci SUS.. Otázky preto obsahovali slovnú zásobu blízku deťom tak, aby otázky dokázali správne zodpovedať. Príkladom zmeny ohľadom slovnej zásoby je napríklad použitie slova "hra", s ktorou sa deti často stretávajú, namiesto slova "systém/aplikácia". Výsledná formulácia otázok pre účastníkov nášho testovania:

1. Chcel(a) by som sa túto hru hrať často.
2. Hra bola ťažká na hranie.
3. Myslím si, že hra bola ľahká na hranie.
4. Potreboval(a) by som pomoc, ak by som sa chcel(a) hrať.
5. Vedel(a) som, čo mám robiť v hre ďalej.
6. Niektoré veci v hre nedávali zmysel.
7. Pre mojich kamarátov by bolo ľahké hrať túto hru.
8. Pri hraní som musel(a) robiť čudné veci.
9. Pri hraní som sa cítil(a) dobre.
10. Pri hraní som sa musel(a) veľa naučiť.

Jednotlivé odpovede boli zaznamenané a na základe pravidiel systému SUS bolo vypočítané výsledné skóre. Nasledujúca tabuľka Tab. 1 obsahuje súčet bodov a vypočítané výsledky SUS.

Účastník	Súčet bodov	SUS skóre
1	27	67,5
2	32	80

Tab. 1 Vyhodnotenie použiteľnosti z pohľadu žiaka

Výsledné SUS skóre aplikácie z pohľadu žiaka je 73,75. To naznačuje dobrú použiteľnosť vytvorenej aplikácie, aj keď stále existuje priestor na zlepšenie.

## VII. ZÁVER

Virtuálna realita a jej technológie predstavujú jednu z najprogressívnejších sa rozvíjajúcich smerov informatiky a informačných technológií v súčasnosti. Najmä webová VR prináša zmenu komunikácie s výpočtovými systémami tak, aby boli pre človeka vrátane ľudí s hendikepom čoraz viac prirodzenejšie, jednoznačnejšie a jednoduchšie s rýchlou odozvou, aj v súvislosti s pandemickou situáciou a využívaním dištančného vzdelávania. V rámci článku bolo predstavené riešenie v podobe vzdelávacej aplikácie na báze zdieľanej webovej VR využiteľnej pri vzdelávaní žiakov so špeciálnymi výchovno-vzdelávacími potrebami. Pretože každé dieťa si zaslúži rovnaké možnosti, či už je hendikepované alebo nie. Naša práca je dôkazom, že technológia virtuálnej reality má potenciál na využitie aj v prostredí slovenského špeciálneho vzdelávania.

## POĎAKOVANIE

Tento príspevok vznikol s podporou projektu KEGA č. 035TUKE-4/2019: „Virtuálno-realistné technológie a vzdelávanie postihnutých ľudí“.

## LITERATÚRA

- [1] Sobota, B.; Hrozek, F.: Virtuálna realita a jej technológie, vol. 1, Košice : TU, 2013, ISBN 978-80-553-1500-3
- [2] Vera, L.; Herrera, G.; Vived, E.: Virtual reality school for children with learning difficulties. In: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology. 2005, s. 338–341
- [3] Sobota, B. - Korečko, Š. - Hudák, M. - Sivý, M.: Svet virtuálno-realistných technológií a zdravotne postihnuté osoby. In: Virtuálno-realistné technológie súčasnosti a vzdelávanie postihnutých ľudí : Actual virtual-reality technologies and handicapped people education. - Košice (Slovensko) : Technická univerzita v Košiciach s. 1-9 . - ISBN 978-80-553-3389-2
- [4] Hudák, M. - Korečko, Š. - Sobota, B.: Advanced User Interaction for Web-based Collaborative Virtual Reality. In: 11th IEEE International Conference on Cognitive Infocommunications : CogInfoCom 2020. - Danvers (USA) : Institute of Electrical and Electronics Engineers s. 343-348 [online, CD-ROM]. - ISBN 978-1-7281-8213-1 (online) - ISSN 2380-7350.
- [5] Sobota, B. - Lodomirjak, M. - Korečko, Š. - Sivý, M. - Hudák, M.: Experimentálna platforma pre prácu s piktogramami pomocou technológií zmiešanej reality. In: Virtuálno-realistné technológie súčasnosti a vzdelávanie postihnutých ľudí = Actual virtual-reality technologies and handicapped people education. - Košice (Slovensko) : Technická univerzita v Košiciach, 2019 s. 10-29 [print]. - ISBN 978-80-553-3389-2
- [6] Shawn L. H., Mcggee L.: Accessibility. URL=<https://www.w3.org/standards/webdesign/accessibility>, 2016
- [7] Nielsen J.: Accessible Design for Users With Disabilities. Nielsen Norman Group. URL=<https://www.nngroup.com/articles/accessible-design-for-users-with-disabilities>, 1993
- [8] Churchill, E., F; Snowdon, D.: Collaborative virtual environments: an introductory review of issues and systems. virtual reality. 1998, roč. 3, č. 1, s. 3–15.
- [9] Dostal, J.: Activating devices and their use in e-Learning–focussed on handicapped students. Procedia-Social and Behavioral Sciences. 2015, roč. 176, s. 284–290.

# Design Of Methodology for Selection Suitable Analytic Platform in the Area of Data Analysis

*Peter MICHALIK, Ladislav POMSAR, Iveta ZOLOTOVA*

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

peter.michalik.2@tuke.sk, ladislav.pomsar@tuke.sk, iveta.zolotova@tuke.sk

**Abstract** – The article identifies key decision criteria when selecting a suitable analytical platform for data analysis. The article also creates a taxonomy of such instruments in the context of a decision-making methodology design. There is a proposed decision-making model which is used in case studies conducted in Slovak business entities. At the same time, the paper also proposes a new method of decision-making criteria valuation based on the methods of experts. The proposed method called ARTICLE is based on the number of occurrences of decision criteria on scientific and business articles because we assume that the authors of such articles are also experts.

**Keywords** – methodology, data analysis, analytic platforms

## I. INTRODUCTION

Data analysis is currently one of the most promising fields of knowledge. Processes involving this spectrum are often complex, and their solutions extend into many other disciplines such as statistics, software engineering, marketing, economics, machine learning, and the like. Data science is, therefore, an interdisciplinary field of processes and systems used to extract potentially useful information, views, or knowledge from various data sources. The author [1] defines data science in a more complex way – a process consisting of 5 main stages - gathering, processing, examining, modeling, and interpreting data. The gathering is the starting point of the data analysis. This is a process where data are grouped either through their own transaction systems or sensors or from other locations such as websites and social networks. Data gathering process includes different issues of a legal nature, thus confirming the interdisciplinary nature of the issue. Data modification is a process of transforming data into a form suitable for their processing via different types of algorithms, machine learning, etc. It is a widespread phenomenon that the acquired data contain incorrect values or even empty fields. Data examination can be described as a real-life application of descriptive statistics since descriptive statistics are often used to get users more familiar with the features of data samples. This process is closely related to data visualization, which is important for better understanding large data volumes. Modeling and interpretation are key steps in many problems because in these phases/ stages are created various predictive models which are nowadays very valuable.

The market is increasingly saturated with analytical tools and technologies that can create predictive models, reports, ad-hoc analysis, data cubes, transactional databases, and many other interrelated artifacts. In this context, there are more frequent problems with deciding what analytical instrument should be chosen. Organizations that have data from different sources want to exploit this potential. However, they often do not know what tool they should use and what type of analysis they need, and sometimes they do not even know what problem they have. In general, this can be for example a sales prediction, clustering customers, targeted marketing, simple reporting, sales analysis, pattern matching inside medical data in order to predict the possibility of a heart attack and many more.

On what basis should an organization decide what it needs and what it wants? Is there any general classification of tools and methods? Is there a taxonomy of parameters that support certain tools? Is there a general methodology on how to decide in case data needs to be analyzed? We believe that it is essential nowadays to address these issues and also are aware of the complexity of the problem. This area is extremely dynamic, and new software services supporting data analysis

come to the market every day, making it difficult to draft such a methodology. However, we are not saying it is impossible. While consistently studying technical and commercial literature, we realized that many algorithms and analysis methods had been used for several decades. It follows that, although the constant data growth poses a problem for their storage, decision-making criteria taxonomy is not as dynamic. Companies still want predictive models, static reports, targeted marketing, etc. The main challenge for data science is to confront its strength with the volume of data, their variability, and accuracy.

The aim is to propose such a methodology that would facilitate the selection of appropriate analytical tools while not being just a classic "benchmarking". Throughout the selection process, a user should get a clear idea of what he needs and what is most important for his data. If the user wants to see only an overview of his sales for each year and for a specific store in a given year, he does not have to install whole data cubes technology. In this case, a spreadsheet with pivot tables would do. Therefore, the methodology output should be a generally defined instrument or module technology, not a specific commercial product.

Creating such a methodology requires a deep analysis of the current state of tools (parameters entering the methodology) that the product contains. There is a high probability that this type of analysis would make use of multi-criteria decision-making, as the parameters of each product will ultimately belong to a certain type of criteria. It is also necessary to create a taxonomy of the parameters that may contribute to the development of theoretical knowledge in the field of expert systems and or decision-making systems in the field of data analysis.

## II. METHODS USED

Given that the selection of an appropriate analytical platform is actually a multi-criteria problem, there is an evident need to use multi-criteria decision-making methods. A subset of these methods is a quantitative valuation of evaluation criteria. Despite the fact that the work weights criteria using experiments, our proposed method ARTICLE is the most similar to that of experts since weights to each criterion are allocated based on the number of occurrences in professional and commercial articles. If we assume that the authors of such articles are experts from the IT sector, we can talk about a modified expert method as specific values of weights, in our case, will be divided proportionally using a linear function in accordance with the number of occurrences in the set of researched articles. More on this subject can be found in the next chapter of the article.

To verify the proposed method, we decided to use a case method because it is an issue that is complex in terms of its verification through quantitative methods in such a short time frame (4 years in our case). One example is a project - Improvement of IT Services through knowledge mining methods with T-Systems Slovakia. In this case, the analysis made use of R-Studio with R programming language. As this and some other projects had been on at a time we have been designing the methodology; it was not possible to select a tool first and then use it. To verify the proposed methodology, we used the reverse approach and therefore deploy an analytical tool in a random way, apply it to the problem, and then through a questionnaire, determine whether the tool meets the objectives and expectations of the customer/person who worked with it. We assume that if a tool does not fall into the same category as a tool that would be the customer's choice, the customer's dissatisfaction would be expressed more accurately using a questionnaire.

## III. METHODOLOGY

This section forms a main part of the article because it processes the methodology proposal as well as all the necessary matters related to the proposal. We would like to point out that although the methodology is designed to select the appropriate category of analytical platforms, the general model (Figure 1) can be applied to the problem stemming from any other domain if the problem tackles a problem choice based on certain criteria.

*Mapping the current state* is essentially the study of technical and commercial literature related to the topic (subsection III-A). It is about a comprehensive analysis and deepening the knowledge in order to make the overview as comprehensive as possible. In the context of our proposed methodology, it is a study of literature in the field of data analysis, business intelligence, tools supporting knowledge discovery, and the development of technologies for processing large data. In addition to the above, it is also necessary to examine already existing methodologies regarding the issue. The first point of the proposed methodology is actually developing the latest trends in the domain - see the first chapter of the thesis.

*The draft decision model* is one of the fundamental theoretical outcomes of the thesis. Indeed, it is necessary to perceive the problem of selecting the appropriate category of analytic tools such as the multi-criteria problem. Section 4.1 states a formal description of a defined problem and three proposed models to choose from, based on the theory of multiple-criteria decision making.

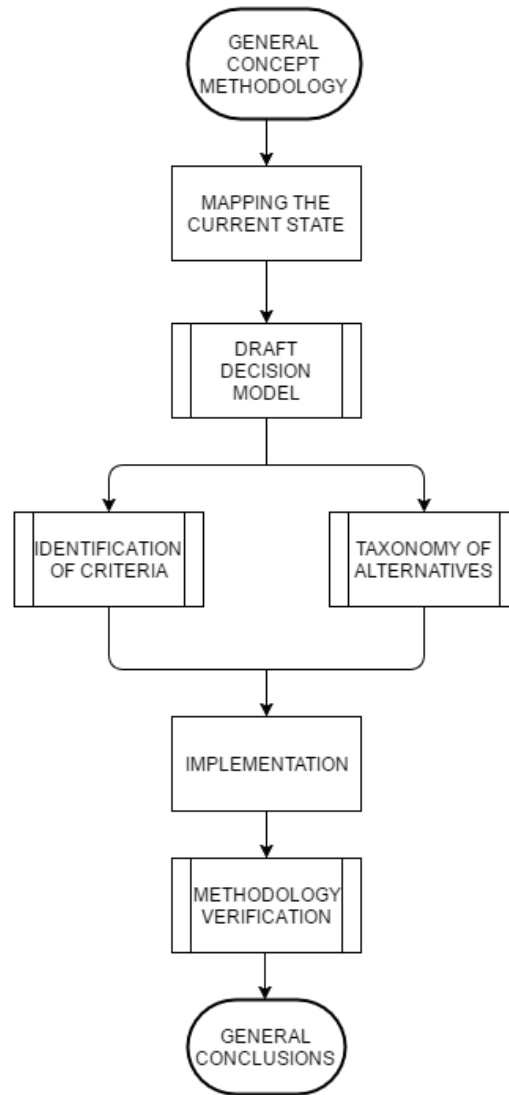


Fig. 1 General concept methodology

*Identification of criteria* is a more time-consuming task. From a set of technical and scientific articles, we identified 34 decision criteria, the description of which is to be found in that paper.

*Taxonomy of alternatives* provides a comprehensive overview of the possible alternatives of the decision-making process and the analytical platform's nature. It also features a description of such alternatives as well as their purpose. Taxonomy of data tools can be perceived as one of the most important theoretical contributions of the work since the present situation is very lively and chaotic. It is often a problem to become familiar with what the current data analysis market actually offers.

*Implementation* in the context of the proposed methodology is mainly about filling the proposed decision-making model with real data obtained in the process of criteria identification and creating a taxonomy of alternatives. The output of this stage is a cross table showing various options and their parameters (criteria).

*Methodology verification* was carried out using case studies. As it is time-consuming and comprehensive, this issue is discussed later. Individual case studies represented real experiments with data provided by companies in Eastern Slovakia.

#### A. Related work

In today's world, there is an increasing number of tools supporting data analysis, with some of them already existing for decades. The first attempts to establish methodologies for their selection and evaluation were available a long time ago - already in the 90s. Even though today it is very difficult to separate data analysis technology into categories, its development has led to its division into three categories - an area of knowledge discovery - Data Mining, support for business analysts

- Business Intelligence and, in recent years, an increasingly large problem with processing large data - Big Data. In many cases, it is easy to identify decision criteria or weights assigned to them, and therefore when designing our own methodology, we rely on multi-criteria decision-making methods. Thus we are speaking about decision-making methodologies and different evaluation procedures of each tool.

The authors [2] [3] undertook to test 14 different DM tools ranging from 75 to 25,000 US dollars. Fascinating is the fact that these instruments were examined by experienced experts in the field as well as students and inexperienced users. This idea is in serious conflict with current trends that prefer experts' views in this area. The methodology consisted of creating 20 different criteria that were then divided into five categories. The scale of the assessment had 6 degrees - excellent, good, average, needs improvement, poor, bad. However, this methodology did not reflect the weights of criteria and so lacked flexibility for decision-making in various areas of data mining. Many research groups are still trying to specify the degree of usefulness of data-mining tools for specific areas such as prediction, detection of fraud, and classification. The authors [4] evaluated five freely available tools designed for prediction models creation – tools that are currently among the market leaders - Weka, RapidMiner, Keel, Orange, and Tangara using 6 selected categories:

- *Graphical interface* – representation, visualization, data management.
- *Input/ Output* – ARF format, other formats, DB collection.
- *Preprocessing* – discretization, handling incorrect values.
- *Types of algorithms* – classification, regression, clustering, association rules.
- *Online/Off-line mode* – online mode, offline mode.
- *Additional functions* – post-processing, meta-learning, statistical tests.

Particularly interesting is the fact that these criteria were assessed in terms of the support provided by the specific technology, divided into four levels - none (N), basic support (B), intermediate support (I), and advanced support (A). If it is impossible to assess the level of support, a binary evaluation is used - yes (Y) or not (N). Moreover, the methodology also featured a simple assessment based on the support for specific types of algorithms.

The authors [5] believe that the identification and evaluation of BI tools must be carried out for the most part on the basis of the most recent data sources (providers, consultancies, scientific and business articles). The remaining part of the analysis should be done in the context of the extant global view of the BI market. The proposed methodology consists of two main parts:

- *Evaluation, classification and selection* of preferred service providers.
- *Final evaluation of selected solutions* based on personal meetings with representatives of the companies providing BI solutions. In this step, the evaluation is easier. Each of the team members evaluates the presentations, and after a mutual agreement, a solution is chosen. The evaluation methodology is essentially freely adjustable according to the internal agreement of the team members.

In addition to the above works, many others deal with criteria identification for selecting suitable analytical platforms from the Big Data domain. The description of these methodologies, however, is not part of this article.

### B. Draft Decision Model - proposed method ARTICLE

The model works with pre-evaluated decision criteria. As mentioned above, the parameters will be selected from the set of research articles. The views of such commercial or scientific articles can be considered experts' views since their authors are the people dealing with this topic on the expert level. Gradually we obtained a finite number of parameters  $n$  entering into this model. The primary added value of this model is the fact that we can allocate a certain weight to these parameters.

Let's assume there is an independent variable  $x_j$  which represents a number of occurrences of a given set of criteria within the set of professional articles. If the criterion is mentioned in an article several times, it is counted as if it was mentioned there only once. Each  $a_j$  parameter has therefore assigned  $x_j$  weight. This means that if the  $a_j$  parameter is the most numerous in the set of 30 articles, it has the highest weight. The question is how to allocate the weight to other parameters and in what proportion. It might be worth trying a number of functional standards. However, we opted for a linear approach. That is, if one criterion occurs 20 times and the second only 10 times, the weight of the second parameter is only half of the first one. The function for weighting will be as follows:

$$y = \frac{1}{n} * x$$

**Table 1** shows an example of weight allocation based on the selected direct proportion function while assuming that we have a finite number of parameters  $n = 24$ .

	$a_1$	$a_2$	$a_3$	$a_4$
x	5	24	3	17
y	0,208	1	0,125	0,708

Table 1  
 Weighting of the input parameters

There is shown that the criterion  $a_1$  has a weight of 0,208 or 20,8 % in relative terms. Analogously, we can explain the other criteria as well: 100 %, 12,5 % and 70,8 %. From now on we will refer to criteria weights as  $w_j$ . Furthermore, we will define the total of all the weights of the criteria featuring in our model:

$$Sw = \sum_{j=1}^n w_j$$

This metric will play a vital role in our proposed algorithm. Until now, all have been seen from the point of view of a user–entity. This means that they cannot in any way affect the functioning of our proposed solution until now. Subsequently, however, the stage comes where an entity has to choose criteria on the basis of which they will be offered a solution in the form of a proposed analytical platform. Again, however, it does not give out its preferences in the orderly form (there is a preference for a selected set of criteria over the others).

The task of an entity is, therefore, to choose the criteria that are attractive for them. However, the question is, how many of those criteria have to be chosen for the given entity? Where is the selection threshold for the given entity? Does this threshold vary depending on the size of the entity? The answer to these questions is complicated to find since the answer will require several iterations and modifications to our proposed model, thus posing a significant obstacle with regard to time. This means we have to opt for an experimental approach, much like in determining weights. It goes without saying that everyone wants his analytical tool to do everything – to identify as many criteria as possible. We must not forget that some of the selected criteria will be in conflict with each other. This fact and also the fact that the price is increasing with each added criterion suggest that the percentage could be more in the bottom half of the spectrum – 50 % or less. Since the methodology has to have a meaning, the percentage cannot be too low. Therefore we decided to set a limit at  $V < 40$  %. It is possible that after a number of iterations and verifications, the percentage will stabilize at some other value.

After we determine the selection limit, the algorithm searches the table with individual attributes. The table may look like the one in the previous model with one difference. In addition to the information on whether the given alternative offers the chosen parameter, the table displays the weight of the given criterion. After going through all the alternatives, we get a value for each alternative stating the number of points (percentage) gained by the specific alternative. The winner is stated similarly as in the previous model - the maximum value of the sum of criteria weights.

$$Sw = \sum_{p=1}^n w_j$$

Such proposed algorithm goes through the following phases:

- 1) Defining the criteria on the basis of a set of scientific articles.
- 2) The allocation of weights to these criteria, based on the occurrence of a given set of articles and the direct proportion function.
- 3) The selection of criteria by the user while the sum of weights of these criteria must not exceed 40% of the total sum.
- 4) Browsing through all the alternatives and quantification of the achieved score for each of the alternatives.
- 5) Stating the winner.

**Table 2** shows the example of a matrix that browses through the algorithm. If the cell is empty, it is clear that the alternative does not fulfill the criterion; for example - it does not support some function.

The total sum of weights = 2.041, 40% of the total weight is 0,82. This means that the algorithm allows users to select only criteria up to this value while regularly informing him about "how many points are left." The example includes a relatively low number of attributes, and thus the resulting sum is also very low. Thus, the user has only two options how to choose the attributes – either he chooses attributes  $a_1$  and  $a_3$  or opts for  $a_4$ . In the former case the winner would

	$w(a_1)$	$w(a_2)$	$w(a_3)$	$w(a_4)$
A1	0,208	1		
A2	0,208	1	0,125	
A3		1	0,125	0,708
A4		0,208		0,708

Table 2  
 The allocation of weights

Criterion	Occurrence	Normalized weight
Data processing in real-time	19	0,0546
Data processing in near real-time	4	0,0115
Batch processing	15	0,0431
Relational data	24	0,0690
Semi-relational data	11	0,0316
Unstructured data	6	0,0172
Integration with RDBMS	24	0,0690
Integration with NoSQL	8	0,0230
Machine Learning	14	0,0402
Integration across technologies	3	0,0086
Existence of GUI	7	0,0201
Integration of visualization libraries	2	0,0057
Data preprocessing	19	0,0546
Integration of programming languages	16	0,0460
Online Analytical Processing - OLAP	21	0,0603
Descriptive statistics	17	0,0489
Data-Marts	3	0,0086
Need to know programming language	11	0,0316
Need to know database language	4	0,0115
Integration with robust computing systems	18	0,0517
Text-mining	6	0,0172
In-Memory	5	0,0144
Stream analytics	9	0,0259
Data visualization	8	0,0230
Process visualization	2	0,0057
Diversity of data sources	8	0,0230
Scorecards	25	0,0718
Dashboards	7	0,0201
Reporting	14	0,0402
Ad-hoc reporting	4	0,0115
Internal ETL mechanism	9	0,0259
Intelligent alerts	1	0,0029
Process mining	2	0,0057
Associative model - Business discovery	2	0,0057

Table 3  
 Criteria identification

be the alternative  $a_2$  as it provides both selected parameters and therefore meets 100 % of the requirements. In the latter case we would have two winners -  $a_3$  and  $a_4$ . However, this is just an example. The real model will feature many more parameters and we can assume that we would see several winners. In the event that this will happen, it is up to the user which of those alternatives he will choose since each of them meets his requirements. We obtain the percentage fulfillment of user's requests from the following relationship:

$$V = 2,5 * \frac{\sum_{j=1}^p w_j}{\sum_{j=1}^n w_j} * 100$$

### C. Criteria identification

Table 3 shows the identified criteria based on the ARTICLE method proposed and explained in the previous chapter of this paper.

The duration of the data processing is an important selection criterion, confirmed by the occurrence of this criterion in the set of articles. In general, there are three types of data processing depending on latency. However, by looking at the values, we can assume that there are only real-time processing and batch processing. Real-time processing gives responses immediately – it is used in financial markets, stock exchanges, various control applications, and the like. On the contrary, batch data processing is a typical example of data warehousing, OLAP analysis, or data from physical measurements [6] [7] [8].

A considerable number of criteria were connected with the nature of the data. We are thus speaking about structured (relational), semi-relational (XML), and unstructured (video, audio) data. Despite the fact that we now face a huge increase in demand for applications that analyze



various social networking sites and a vast amount of unstructured data, it is clear that spreadsheet format data still play a major role in the issue. This issue is indirectly related to the possibility of integration with relational database management systems (RDBMS) and a variety of NoSQL databases. Integration with other types of databases (NoSQL) was not recorded in the selected set of articles [6] [9].

The existence of libraries for machine learning is also very interesting. However, the problem here is that this group covers an enormous number of algorithms, and we face the question of whether it will be possible to divide this massive group into a number of smaller units, thus creating several simpler criteria. The truth is that this criterion, in most cases occurred in the standard term - support for machine learning algorithms [10] [11].

Integration across technologies is a criterion that appears to be vague. Nevertheless, it is essential to take this criterion into account to avoid a situation where it is impossible to merge, combine and integrate the analytical platform with other technology from another supplier, etc. [12] [13] [14].

The existence of a GUI is one of the oldest software technology requirements. This reflects that it is difficult for many people to work only in one programming mode, and working with GUI becomes more interesting and more attractive in their eyes [14].

A wider range of decision-making criteria includes visualization and related services. In principle, the visualization, in this case is a part of an analytical tool, and now it can be taken for granted (as evidenced by its relatively high number of incidence). Then there is the criterion of process visualization, which is likely to be related to data mining processes. However, its occurrence is negligible. The third group is tools for visualization support which can be integrated into one's own system. These are, in fact, different libraries that often become a part of different web applications while also serving for data analysis.

Preprocessing is one of the critical moments of data analysis, as evidenced by the number of occurrences. Preprocessing is made up of different transformation processes, values editing, statistical methods for solving problems with incomplete data, etc. Basically, it is the second phase of the ETL process, which is also one of the required parameters of analytical platforms. It also provides the opportunity to work with a variety of heterogeneous sources such as CSV, XLXS, or XML. [8]

Programming languages are yet another category posing several problems. There is a need for different (or at least one of the selected) programming languages for better data manipulation. In contrast, some people avoid programming. In this respect, the following parameters have been identified - the need to know programming (Python, Java) and database languages (SQL).

Integration within a robust computing system is also one of the criteria. For instance, we are talking about the possible integration with the Hadoop system, which is currently widely used to analyze vast amounts of data. However, the system operates in the dispensing mode and therefore is not suitable for algorithms that operate on the number of iterations. Many experts claim that this system is not suitable for real-time data processing. Therefore there is a need for yet another criterion - support for processing data streams (essentially, it is almost a duplication of the real-time processing criterion). However, we assume that we are talking about the integration with tools such as Apache Spark. [15][16]

The largest number of criteria is criteria related to the analysis of historical data. We are speaking about traditional BI platforms such as support of OLAP analysis, data marketplaces, various kinds of reporting, or scorecards. These tools are typical examples of batch data processing, and it can therefore be assumed that even today, there is still a demand for particular analytical tools that processes mainly historical data. Therefore the lowest possible latency is not a total necessity in the field of Business Intelligence. Descriptive statistics is a tool that gives us an excellent picture of the initial data. However, it is questionable whether a separate criterion suits these characteristics since many other mentioned tools offer descriptive statistics.

Another criteria group is support for text-mining - currently highly preferred. Thus we speak of the existence of algorithms that are part of machine learning and also support the analysis of semi-structured data since these data come mainly from social networks in the form of XML.

An interesting issue is the support of in-memory technology where data are stored in RAM - this reduces the response to data demand by several times. This technology is mainly used by the QlikView platform, an important representative of the associative data model - Business Discovery [17].

The last criterion is the existence of process mining tools - processes taking place in an entity are identified in reverse based on semi-structured data (especially weblogs, server logs, etc.). This problem is closely related to the BPM area and therefore is a bit more specific. It is an essential part of operations taking place in entities where business processes are not clearly mapped. Improved processes, in this case, can streamline business processes, resulting in improved performance and

thus increased profits. It is assumed that the modules designed for finding processes in data records will be a key area of data analysis in the near future, even though the number of occurrences for this parameter was very low.

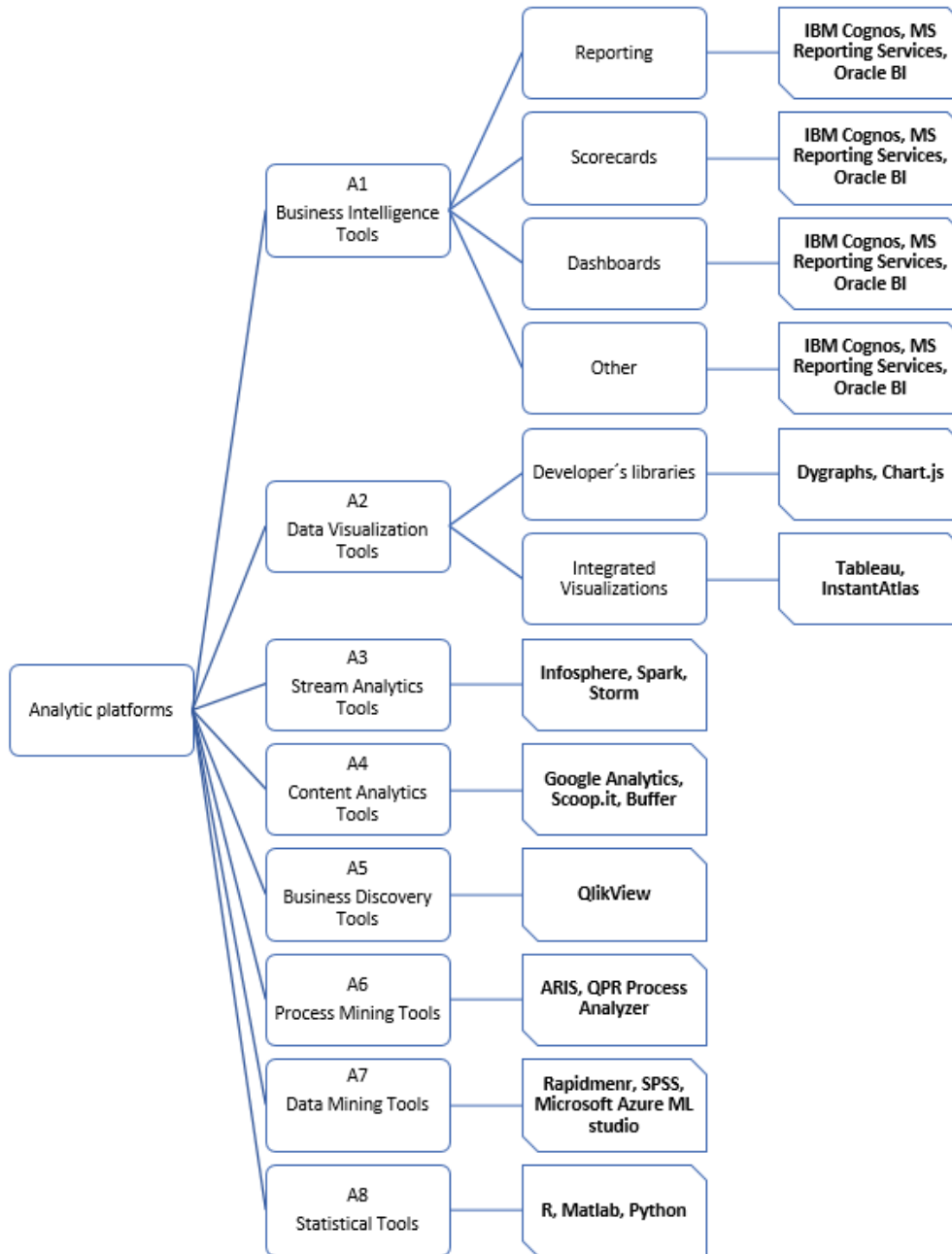


Fig. 2 Tool taxonomy

*D. Tool taxonomy*

Figure 2 shows the proposed taxonomy of analytical platforms. Items A1 - A8 show available alternatives within the proposed methodology for choosing the appropriate analytical tool/ category of tools. Each of these tools is special due to its different decision-making criteria. Specific solutions are marked in bold.

Currently, it is not easy to accurately categorize tools because many suppliers of such solutions are trying to cover the greatest possible set of analyses with their tools. The market, therefore, features complex tools that "can do it all". Therefore, it is possible that, as in the case of BI tools, there are alternatives that offer content analysis. However, this taxonomy aimed to categorize tools

based on their typical features, and therefore BI is a classical tool for tracking historical data [18] [19] [20].

### E. Implementation and Case studies

The starting point of the whole process was to establish contact with an actual business entity that has a data feed from a field it operates in. After obtaining a set of data, we chose the analytical platform independent from the nature and amount of data – Table 4. Subsequently, we deployed the analytical platform to a set of data with a specific purpose, which had been consulted in the company.

Code of case study	Analytical tool	Product	Domain
001	Data Mining Tools	MS SQL Server 2012 - Data Mining Module	Data from the production line - motor assembly
002	Business Intelligence Tools	MS SQL Server 2012 - Reporting Services, Report Builder 3.0	Data from the production line - motor assembly
003	Data Mining Tools	Rapidminer	Data from monitoring of IT systems environment
004	Statistical Tools	R	Data from monitoring of IT systems environment
005	Statistical Tools	Matlab	Data from Configuration Management Tool
006	Business Intelligence Tools	IBM Cognos	Data from the production line – Iron processing
007	Business Intelligence Tools	Excel	Data from process orders in a manufacturing company
008	Business Discovery Tools	QlikView	Meteorological data

Table 4  
List of case studies

The data analysis made use of analytical tools; however, not in line with the recommended methodology. There is a presumption that if a tool was not a right choice (not compliant with the recommended methodology), feedback received from the questionnaire would also point out the tool's inadequacy. In this case, the methodology can be considered appropriate. However, there is also a possibility that the appropriateness of the methodology is not sufficient, and another iteration process in the identification criteria will have to take place. ?? shows the same samples from case studies for illustration only. Table 5 presents the achieved results.

Code of case study	FUR - methodology	FUR - questionnaire	FUR - best alternative
001	A7 - 58,91%	A7 - 40%	A5, A1 - 75,43%
002	A1 - 75,43%	A1 - 90%	A5, A1 - 75,43%
003	A7 - 75,43%	A7 - 60%	A1 - 79,02%
004	A8 - 63,94%	A8 - 60%	A1 - 79,02%
005	A8 - 69,94%	A8 - 50%	A1 - 87,64%
006	A1 - 72,56%	A1 - 100%	A1 - 72,56%
007	A1 - 86,21%	A1 - 100%	A1 - 86,21%
008	A5 - 99,14%	A5 - 80%	A5, A1 - 99,14%

Table 5  
Results

Case studies 006 and 007 are very suitable for interpretation of results. Designed methodology recommends tools from category A1 for data analysis and also used tool is from that same category. It means that fulfillment of user's request (FUR) according questionnaire should be also very high. It's true in this case and it confirms correctness of proposed methodology as well. On the other side, results from case study 003 are very questionable because the fulfillment of user's request according to the questionnaire is relatively poor (60 %) considering the fact that the A7 alternative was the second-best option with 75,43%. The best alternative was A1 with 79,02%.

## IV. CONCLUSION AND FUTURE WORK

A major conclusion from the case studies is that the identified criteria are meaningful, and their assigned weights have their own meaning. Also, the weight limit set at 40 % appears to be acceptable. However, we do not rule out the possibility that the dynamic development of the area may force future researchers to update the methodology by adding or removing other criteria. We believe, however, that the way these criteria are identified and chosen may remain the same.

## ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-20-0247.

## REFERENCES

- [1] J. Janssens, *Data Science at the Command Line: Facing the Future with Time-Tested Tools*. " O'Reilly Media, Inc.", 2014.
- [2] M. A. King, J. Elder, B. Gomolka, E. Schmidt, M. Summers, and K. Toop, "Evaluation of fourteen desktop data mining tools," in *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)*, vol. 3. IEEE, 1998, pp. 2927–2932.
- [3] K. Collier, B. Carey, D. Sautter, and C. Marjanemi, "A methodology for evaluating and selecting data mining software," in *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*. IEEE, 1999, pp. 11–pp.
- [4] N. Hasim and N. A. Haris, "A study of open-source data mining tools for forecasting," in *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*, 2015, pp. 1–4.
- [5] M. Mircea *et al.*, "Strategy for selecting a business intelligence solution," *Revista Informatica Economică nr.*, vol. 1, no. 45, p. 103, 2008.
- [6] D. van Beek, "Selecting BI tools," May 2021, [Online; accessed 30. May 2021]. [Online]. Available: <https://www.passionned.com/bi/tools/choosing-and-selecting-bi-tools>
- [7] "Business Intelligence (BI) Tool Selection Process," May 2021, [Online; accessed 2. February 2021]. [Online]. Available: <https://www.rpi.edu/datawarehouse/docs/BI-Tool-Selection.pdf>
- [8] "How To Choose The Best Tool For Your Big Data Project - Big Data Analytics News," Apr 2014, [Online; accessed 30. May 2021]. [Online]. Available: <https://bigdataanalyticsnews.com/choose-best-tool-big-data-project>
- [9] "Selecting Data Mining Tools to Increase Customer Value," Jan 2004, [Online; accessed 30. May 2021]. [Online]. Available: <https://www.destinationcrm.com/Articles/Web-Exclusives/Viewpoints/Selecting-Data-Mining-Tools-to-Increase-Customer-Value-44404.aspx>
- [10] J. Ranjan and V. Bhatnagar, "A review of data mining tools in customer relationship management," *Journal of Knowledge Management Practice*, vol. 9, no. 1, 2008.
- [11] D. Crolene, "Business Intelligence Tools," May 2021, [Online; accessed 30. May 2021]. [Online]. Available: <http://ds.datasourcesconsulting.com/blog/business-intelligence-tools>
- [12] X. L. Dong and D. Srivastava, "Big data integration," in *2013 IEEE 29th international conference on data engineering (ICDE)*. IEEE, 2013, pp. 1245–1248.
- [13] T. Lu, X. Guo, B. Xu, L. Zhao, Y. Peng, and H. Yang, "Next big thing in big data: the security of the ict supply chain," in *2013 International Conference on Social Computing*. IEEE, 2013, pp. 1066–1073.
- [14] E. Begoli and J. Horey, "Design principles for effective knowledge discovery from big data," in *2012 Joint Working IEEE/IFIP Conference on Software Architecture and European Conference on Software Architecture*. IEEE, 2012, pp. 215–218.
- [15] R. Schneider, *Custom hadoop for dummies, special edition*. Place of publication not identified: John Wiley & Sons Inc, 2012.
- [16] G. Turkington, *Hadoop Beginner's Guide*. Packt Publishing Ltd, 2013.
- [17] P. Hajastek, P. Michalik, and I. Zolotova, "Analýza meteorologických dát s využitím nástroja qlikview," *Electrical Engineering and Informatics*, vol. 5, p. 896–900, 2014.
- [18] P. Michalik, J. Štofa, and I. Zolotova, "Concept definition for big data architecture in the education system," in *2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. IEEE, 2014, pp. 331–334.
- [19] V. Hahanov, W. Gharibi, E. Litvinova, and S. Chumachenko, "Big data driven cyber analytic system," in *2015 IEEE International Congress on Big Data*. IEEE, 2015, pp. 615–622.
- [20] C.-W. Wang and S. E. Chang, "Big data in practice: A study on cloud data collection and storage services." in *ICADIWT*, 2016, pp. 211–219.

# Tvorba zabudovaných systémov snímačov pre potreby inteligentného priestoru

Ján VAŠČÁK

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

jan.vascak@tuke.sk

**Abstrakt** – Hlavným cieľom tohto článku je navrhnúť a implementovať rozšírenia pre už existujúci inteligentný priestor na domácom pracovisku pomocou vhodného hardvéru a softvéru. Hardvérové prvky zahŕňajú rôzne snímače a akčné členy a nové softvérové prvky predstavujú vytváranie súkromného cloudu. V článku je uvedený teoretický prehľad použitých technológií, po ktorých nasleduje nový návrh a implementačný popis rozšírenia pôvodného inteligentného priestoru. Konkrétnejšie boli vytvorené štyri nové zariadenia, ktoré obsahujú rôzne snímače a akčné členy, ktoré sú schopné plne spolupracovať so súkromným cloudom, ktorý bol špeciálne navrhnutý pre tento účel.

**KLúčové slová** – cloud, inteligentný priestor, internet vecí, snímače

## I. ÚVOD

Pojem inteligentný priestor (IS) nie je jednoznačný. Pokrýva mnoho hardvérových a softvérových zdrojov, ktoré sa používajú na digitalizáciu tohto priestoru. Pomocou týchto zdrojov sme schopní rýchlo a ľahko získať prístup k informáciám, ktoré človek ináč sám nedokáže vnímať. Takýto priestor je zvyčajne určený na zabezpečenie interakcie s jeho používateľmi, ale jeho vytvorenie a implementácia nie sú triviálne.

Tento príspevok sa zaoberá softvérovým a hardvérovým rozšírením pre už existujúci IS na domácom pracovisku Katedry kybernetiky a umelej inteligencie (Technická univerzita v Košiciach), ktorého štruktúrny a funkčný popis nájdete v [1], [2]. Na začiatku bola potreba vyriešiť potreby vlastného cloudového výpočtového prostredia. V súčasnosti existuje veľa poskytovateľov cloudových služieb tretích strán, ale ich riešenia zvyčajne nie sú až tak lacné a používatelia sú viazaní na jednu technológiu, takže je takmer nemožné migrovať medzi poskytovateľmi cloudov. Existuje tiež riziko, že poskytovateľ cloudových služieb ukončí toto podnikanie, čo nám ako koncovému používateľovi spôsobí veľa problémov.

V tejto práci sú tieto problémy vyriešené vytvorením vlastného cloudového riešenia s vlastným hostingom a s ľahko použiteľným webovým rozhraním, ktoré sa používa na správu zdrojov používateľov. Bola postavená na existujúcom hardvéri, takže nevznikli žiadne ďalšie náklady. Taktiež bolo potrebné rozšíriť existujúci IS o nové hardvérové možnosti, ako sú nové snímače a akčné členy. Tieto potreby boli splnené vybudovaním vlastných IoT modulov, ktoré boli zostavené pomocou Raspberry Pi a súprav snímačov. Ich funkčnosť bola vyhodnotená pomocou vlastného hostingového cloudového riešenia.

V ďalších častiach si popíšeme niekoľko základných pojmov, ktoré sú potrebné pre pochopenie prezentovaného problému, a to IS, internet vecí (IoT) a cloud computing. Po krátkom prehľade požiadaviek na infraštruktúru nasleduje návrh na rozšírenie už existujúceho IS. Potenciál takého rozšírenia bude následne ozrejmený na skutočnej aplikácii a niektoré záverečné poznámky tento článok završujú.

## II. ZÁKLADNÉ ZÁSADY V BUDOVANÍ INTELIGENTNÉHO PRIESTORU

Pojem *inteligentný priestor* popisuje miesto, kde je umiestnených a vzájomne komunikujúcich veľké množstvo snímačov a inteligentných zariadení (väčšinou IoT). Za IS môžeme považovať izbu, halu, budovu, alebo dokonca ulicu vybavenú snímačmi a akčnými členmi. Pomocou snímačov zhromaždíme údaje o tom, čo ľudia v tomto priestore robia, kde sa ľudia v tomto priestore nachádzajú a pomocou akčných prvkov s nimi môžeme interagovať. IS je priestor, kde môžu

ľudia a počítače navzájom komunikovať a kde môžu roboty spolupracovať s ľuďmi. Hlavnou myšlienkou IS je distribúcia snímačov a akčných členov v priestore s cieľom znížiť potrebu výpočtového výkonu pre roboty, úspešne jestvovať a pohybovať sa v takom priestore [3].

IoT je nový pojem v oblasti telekomunikácií, ktorý definuje, že všetky objekty každodenného života budú v budúcnosti vybavené mikroprocesorom schopným komunikovať cez internet pomocou vhodných protokolov. Aj tieto veci budú komunikovať s používateľmi a stanú sa neoddeliteľnou súčasťou každodenného života. Hlavným cieľom IoT je zaručiť ľahký prístup a interakciu so širokou škálou zariadení, ako sú domáce spotrebiče, bezpečnostné kamery, vozidlá atď. [3].

Cloud computing je výpočtový systém na požiadanie (on demand). Pokrýva hlavne ukladanie dát a výpočtové zdroje. Tento termín sa všeobecne používa na označenie dátových centier, ktoré sú dostupné mnohým používateľom prostredníctvom internetu. Veľké cloudy často distribuovali svoje služby na širokú škálu miest. Ak je cloud relatívne blízko používateľa, možno ho považovať za hraničný výpočtový systém.

Cloud môže byť používaný jednou organizáciou (súkromný, či privátny cloud), verejne dostupným (verejný cloud), alebo kombináciou oboch prístupov (hybridný cloud). Najväčším poskytovateľom verejného cloudu v súčasnosti je Amazon so službou Amazon Web Services (AWS). Jednou z najväčších výhod cloud computingu je jeho schopnosť škálovať zdroje na základe dopytu a znížiť náklady na údržbu hardvéru používateľov. Používatelia sú zodpovední iba za svoje aplikácie, ich nasadenie v cloudovej službe a za konfiguráciu potrebných zdrojov.

Vývoj IoT do značnej miery závisí od vysokého výpočtového výkonu a vysokej dostupnosti zdrojov, čo je hlavná výhoda služieb cloud computingu. To je hlavný dôvod, prečo je cloud computing základnou potrebou IoT. Cloudové výpočty poskytujú tri typy služieb [4]:

- infraštruktúra ako služba (IaaS),
- softvér ako služba (SaaS),
- platforma ako služba (PaaS).

*Infraštruktúra ako služba* poskytuje výpočtové zdroje (CPU, RAM, atď.), ukladanie dát a sieťové možnosti. Spravidla obsahuje orchestrátor ako Open Stack alebo Apache Cloudstack. Tento orchestrátor sa používa na správu a vytváranie virtuálnych strojov a rozhoduje, ktorý fyzický stroj mu bude zabezpečovať hosting. Umožňuje tiež migráciu virtuálnych strojov medzi hosťiteľmi. Hlavnou alternatívou virtualizácie je kontajnerizácia, ktorá umožňuje spúšťanie kontajnerov pomocou izolovaných častí s jedným jadrom Linuxu, ktoré má priamy prístup k hardvéru. Jadro Linuxu obsahuje veľa technológií, ktoré sa používajú na izoláciu a zabezpečenie týchto kontajnerov. Kontajnerizácia ponúka vyšší výkon ako virtualizácia s menšou réžiou a je ľahko škálovateľná na základe využitia [5].

*Softvér ako služba* je počítačový program, alebo súbor počítačových programov, ktorý sa používa ako hosťiteľ aplikácií, ktoré sú v ňom poskytované. SaaS možno v zásade považovať za typ špecializovaného aplikačného servera. Vývojári softvéru, ktorí používajú softvér ako službu, môžu vyvíjať svoje aplikácie pre jedného zákazníka a potom distribuovať svoj softvér pomocou SaaS viacerým zákazníkom s takmer nulovým dodatočným nákladom. Táto platforma pokrýva správu všetkých potrebných hardvérových a softvérových zdrojov a tieto zdroje využíva na poskytovanie aplikácií pre zákazníkov na požiadanie. Tento prístup sa používa na skrátenie času na vývoj vysoko dostupných podnikových aplikácií [6].

Platforma ako služba dáva do popredia produktivitu vývojárov, čo jej umožňuje vytvoriť vrstvu abstrakcie nad IaaS pre rýchle a ľahké nasadenie služieb. PaaS sa primárne zameriava na produktivitu vývojárov, potom na správu siete, ukladanie dát a výpočtový výkon. Vďaka tomu je schopný abstrahovať implementáciu mimo runtime aplikácie. Vývojárom je poskytované prostredie a služby, ktoré sa používajú na spustenie programovacieho jazyka a infraštruktúra je ponechaná na nižšie vrstvy ako napr. IaaS [7].

### III. VŠEOBECNÉ POŽIADAVKY NA INFRAŠTRUKTÚRU

Pre správnu implementáciu IS je potrebné splniť celú radu požiadaviek ako napríklad:

- snímače / akčné členy,
- softvér,
- ostatný hardvér.

Snímače sa používajú na poskytovanie informácií o danom prostredí v reálnom čase. Môžu byť použité napríklad na detekciu množstva denného svetla, hladiny hluku, teploty, vlhkosti alebo môžu byť použité v robotoch na detekciu prekážok a terénu. Snímače sú v zásade rozdelené do dvoch hlavných kategórií, a to aktívne a pasívne snímače.

Aktívne snímače emitujú určitý druh energie z hľadiska merania. Napríklad ultrazvukové snímače na detekciu prekážok vysielať ultrazvukovú energiu, ktorá sa odráža od povrchu a naspäť sa zachytáva. Pomocou známych hodnôt rýchlosti zvuku môžeme vypočítať vzdialenosť objektu.

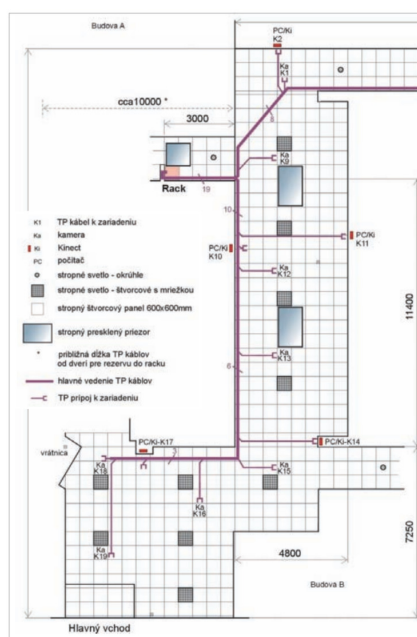
Pasívne snímače zase prijímajú určitý druh energie a na základe jej intenzity môžu vypočítať niektoré užitočné údaje.

Aby boli všetky snímače plne funkčné, potrebujú nejaký softvér, ktorý im umožňuje vzájomnú komunikáciu a ukládanie niektorých častí údajov. Cloud computing sa v tomto prípade javí ako jedno z najlepších riešení. Na vytvorenie vlastného cloudového výpočtu je potrebný určitý druh virtualizácie alebo kontajnerizácie a určité rozhranie, aby s ním používatelia mohli pracovať.

Taktiež je potrebný ďalší hardvér. Prvý z nich je riadiaci modul pre snímače a akčné členy. Jeho úlohou je poskytovať im pripojenie na internet a musí byť schopný ich ovládať a zhromažďovať ich údaje. Na spracovanie údajov a cloud computing je potrebný nejaký fyzický počítačový hardvér známy ako server. Tento server musí byť hosťiteľom zákazníckych cloudových riešení a poskytovať výpočtové možnosti pre radiče, ktoré sú vybavené snímačmi. Určitý druh predspracovania a filtrovania je možné vykonať priamo na týchto radičoch, ale čím sú náročnejšie úlohy, tým je viac potrebný výkonnejší hardvér.

#### IV. NAVRHNUTÉ ROZŠÍRENIE INTELIGENTNÉHO PRIESTORU

Hlavným účelom tejto práce bolo rozšírenie existujúcich IS. V súčasnosti sa už používa veľa inteligentných zariadení, ako sú fotoaparáty, 3D skenery, alebo mikrofóny. Všetky sú pripojené k mini PC, ktoré sa pre ne používa ako ovládač a sú pripojené k lokálnej sieti, vid' Obr. 1.



Obr. 1 Súčasný stav IS a jeho prepojenia.

Najskôr bolo potrebné implementovať nové snímače a akčné členy. Hlavným cieľom bolo vytvoriť štyri sady IoT zariadení s rôznymi snímačmi.

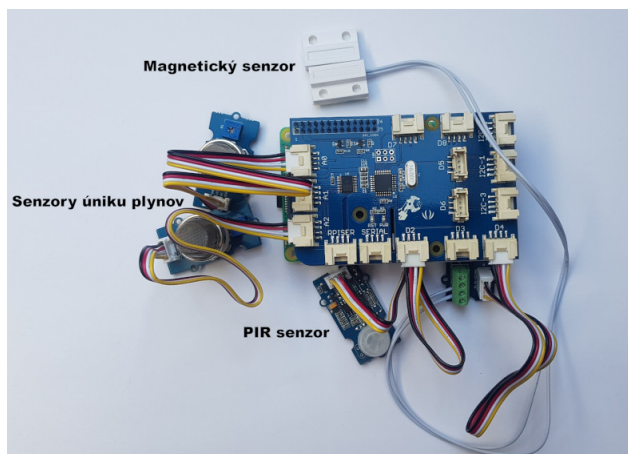
Z hľadiska potreby vytvorenia súprav IoT boli použité štyri rôzne sady snímačov a akčných členov. Ako radič bol použitý Raspberry Pi s nainštalovaným operačným systémom Raspbian. Tieto ovládače poskytujú internetové pripojenie k týmto snímačom a poskytujú užitočné knižnice potrebné na komunikáciu so snímačmi a akčnými členmi.

Prvou z týchto súprav je *sada bezpečnostných snímačov*. Skladá sa zo snímačov používaných hlavne na zaistenie bezpečnosti, ako je magnetický snímač dverí, PIR senzor, snímač úniku plynu a ďalšie, vid' Obr. 2.

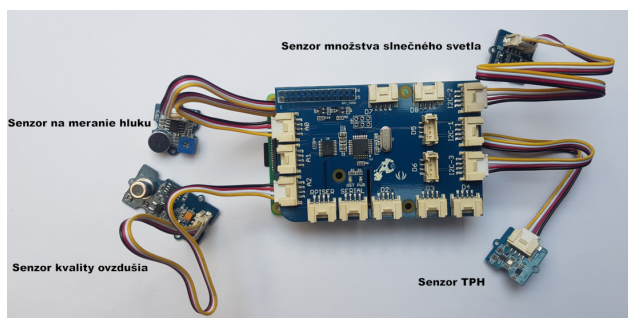
Ďalšou súpravou snímačov je *sada snímačov životného prostredia*. Jej hlavným účelom je poskytnúť informácie o životnom prostredí, ako je teplota, vlhkosť, tlak, množstvo slnečného žiarenia, atď., vid' Obr. 3.

Tretím z nich je *sada lokalizačných a navigačných snímačov*. Táto sada je užitočná na lokalizáciu a navigáciu robotov v IS. Poskytuje informácie o vzdialenosti najbližších objektov, orientácii robotov, alebo dokáže detegovať bielu čiaru na zemi, vid' Obr. 4.

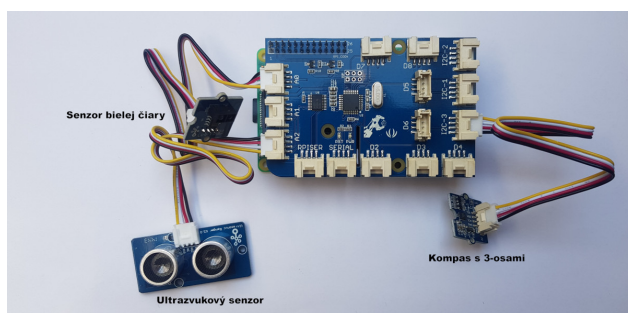
Poslednou použitou súpravou je *základná sada snímačov*. Používateľom má poskytnúť základné snímače a akčné členy, ktoré im pomôžu pochopiť, ako pracovať s pokročilejšími snímačmi. Obsahuje snímače ako tlačidlá, alebo svetelné snímače a tiež akčné členy ako LED, bzučiak, alebo OLED displej, vid' Obr. 5.



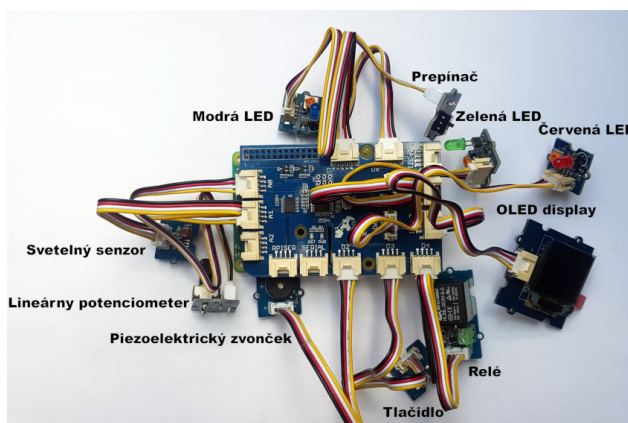
Obr. 2 Zobrazenie sady bezpečnostných snímačov.



Obr. 3 Zobrazenie sady snímačov životného prostredia.



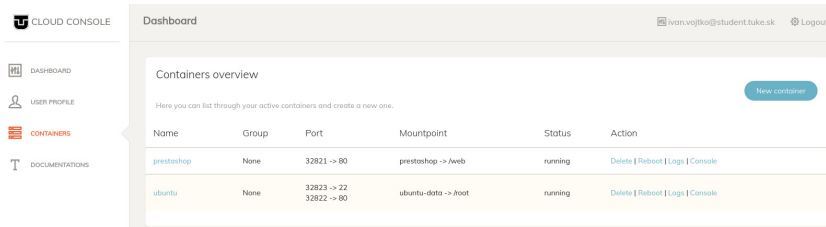
Obr. 4 Zobrazenie sady lokalizačných a navigačných snímačov.



Obr. 5 Zobrazenie základnej sady snímačov.



Aby súpravy IoT mohli medzi sebou kooperovať, bolo vyvinuté vlastné riešenie cloud computingu. Hlavnou technológiou, ktorá pomáha tomuto cloudu pracovať podľa ich potreby, je *Docker*. Jedná sa o kontajnerizačnú službu, ktorá má jednoduché nasadenie riešení a poskytuje používateľom mnoho vopred pripravených obrazov potrebných na spustenie rôznych typov aplikácií, programovacích jazykov a prostredí. Toto riešenie je riadené a spravované webovou aplikáciou Flask, ktorá je napísaná v jazyku Python a poskytuje užívateľom ľahko použiteľné riešenia. Používatelia sú schopní ľahko vytvárať a spravovať svoje kontajnery a nasadiť na nich aplikácie. Používateľom tiež poskytuje v prehliadači terminál pripojený k týmto kontajnerom, vid' Obr. 6.



Obr. 6 Cloudovské rozhranie so spustenými kontajnermi.

Toto cloudové riešenie má tiež k dispozícii API na automatizáciu procesu nasadenia a dokumentácie, ktoré obsahuje informácie o všetkých hardvérových a softvérových prostriedkoch dostupných v tomto IS.

## V. PRÍKLAD APLIKÁCIE REÁLNEHO ŽIVOTA

Tento vylepšený IS je schopný poskytnúť celú radu možností pre experimenty, napr. inteligentná klimatizácia. Pomocou cloudu môžeme nasadiť dve čiastkové aplikácie. Jedna so spusteným brokerom MQTT a druhá s databázou, napríklad MongoDB. Potom môžeme pomocou bezpečnostnej súpravy IoT monitorovať, či sa v tomto priestore nachádzajú nejaké osoby a pomocou súpravy snímačov životného prostredia dokážeme monitorovať priame slnečné svetlo, teplotu, vlhkosť, kvalitu vzduchu a hlásiť tieto informácie prostredníctvom cloudu do databázy. Potom základná sada snímačov na základe poskytnutých informácií rozhodne, či je potrebné klimatizáciu zapnúť alebo vypnúť.

Taktiež môže byť do existujúceho robota implementovaná sada lokalizačných a navigačných snímačov, ktorá mu môže poskytnúť ďalšie snímače. Dáta zhromaždené snímačmi sa môžu posielat' do cloudu, ktorý môže rozhodnúť, čo robiť a ako napláňovať pohyb a cestu robota.

V tomto IS boli vykonané jednoduché experimenty. Prvým bolo sledovanie prítomnosti osôb vo vnútri tohto priestoru. Použila sa sada bezpečnostných snímačov a základná sada snímačov. Tento experiment využíva magnetický dverový snímač na detekciu, či niekto vstupuje do miestnosti, alebo z nej odchádza a PIR senzor na elimináciu chýb, ktoré by mohli byť spôsobené otvorením, resp. zatvorením dverí vetrom.

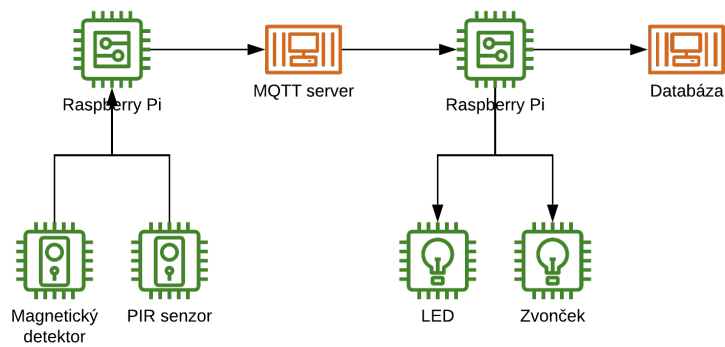
Súprava bezpečnostných snímačov monitorovala dvere a prítomnosť osoby, ktorá odosiela tieto informácie prostredníctvom MQTT nasadeného v cloudu do základnej súpravy snímačov, ktorá následne rozsvietila LED diódu, keď boli dvere otvorené a zapla bzučiak, keď bola zistená osoba. Všetky tieto informácie boli potom odoslané do databázy, vid' Obr. 7. Tento jednoduchý experimentálny príklad nám poskytol informácie o tom, ako môže celý tento IS kooperovať a aké úspešné bolo jeho rozšírenie.

## VI. ZÁVER

V tomto článku sme popísali základné myšlienky IS a IoT, a to, čo bolo v súčasnosti k dispozícii v našom IS na univerzite. Úspešne sme rozšírili potenciál tohto IS o nové snímače, akčné členy, ovládače a vlastné cloudové výpočtové riešenie. Tieto prostriedky boli tiež použité v niekoľkých experimentoch, kde sa ukázalo, že implementácia týchto riešení bola úspešná a poskytla nám cenné rozšírenie. V budúcnosti sa tento priestor ešte viac rozšíri o viac snímačov a tieto snímače môžu využívať nasadené cloudové riešenia na vylepšenie svojej funkčnosti.

## LITERATÚRA

- [1] D. Čurová, R. Haluška, T. Hucec, M. Puheim, J. Vaščák, and P. Sinčák, "Intelligent space at center for intelligent technologies — system proposal," in *The 15th IEEE International Symposium on Applied Machine Intelligence and Informatics (SAMII)*, Jan 2017, pp. 000 191–000 196.
- [2] J. Vaščák, E. Kajátí, and I. Zolotová, "Concept of intelligent space in education of IoT applications in robotics," in *The 16th IEEE International Conference on Emerging eLearning Technologies and Applications (ICETA)*, Nov 2018, pp. 629–634.



Obr. 7 Príklad aplikácie bezpečnostnej schémy.

- [3] I. Vojtko, "Sensory Systems Based on RFID Technology in Intelligent Space (in Slovak)," Master's thesis, Technical University of Košice, Slovakia, 2017.
- [4] J. Rui and S. Danpeng, "Architecture design of the internet of things based on cloud computing," in *The IEEE Seventh International Conference on Measuring Technology and Mechatronics Automation*, June 2015, pp. 206–209.
- [5] A. Amies, H. Sluiman, Q. Tong, and G. Liu, *Developing and Hosting Applications on the Cloud*, ser. IBM Press. Pearson Education, 2012.
- [6] G. Liu, "Research on independent SaaS platform," in *The 2nd IEEE International Conference on Information Management and Engineering*, April 2010, pp. 110–113.
- [7] R. Dua, A. R. Raja, and D. Kakadia, "Virtualization vs containerization to support PaaS," in *The IEEE International Conference on Cloud Engineering*, March 2014, pp. 610–614.

# Matematický model výkonového multiport meniča

<sup>1</sup>Adrián MARCINEK, <sup>2</sup>František ĎUROVSKÝ

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

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

<sup>1</sup>adrian.marcinek@tuke.sk, <sup>2</sup>frantisek.durovsky@tuke.sk

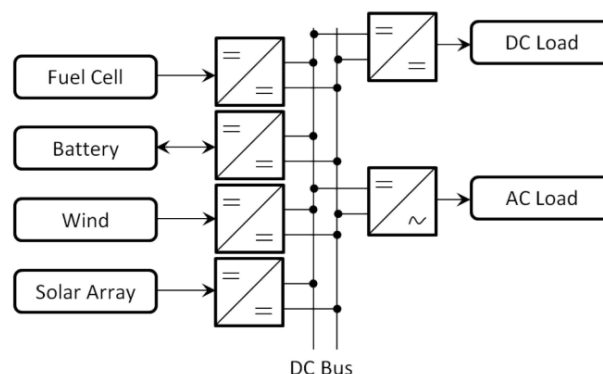
**Abstrakt** — Multiport výkonové meniče sa používajú na pripojenie rôznych druhov zdrojov energie a záťaží. Tieto zdroje sú zvyčajne pripojené pomocou DC-DC meničov na spoločnú jednosmernú zbernicu. Takýto prístup však vyžaduje viacnásobné konverzie napájania, čo znižuje účinnosť celého systému. Článok popisuje matematický model multiport meniča s paralelným pripojením transformátorov. Navrhovaná topológia využíva menší počet výkonových konverzií, čo zvyšuje účinnosť celého systému. Topológia ďalej umožňuje ľahké pridanie alebo odobratie ďalších zdrojov alebo záťaží. Nový prístup sa tiež používa pri odvodzovaní matematického modelu prevodníka.

**Kľúčové slová** — obnoviteľný zdroj energie, elektromobil, výkonový menič, modulárny multiport výkonový menič

## I. ÚVOD

V súčasnosti sa významná časť emisií produkuje v priemysle, vykurovaní budov a tiež v celosvetovej doprave, najmä v cestnej doprave. Existuje niekoľko spôsobov, ako je možné znížiť emisie. Jedným z nich je využitie obnoviteľných zdrojov energie, ako sú palivové články, fotovoltika, batériové systémy atď. Primárnym dopytom po týchto systémoch je spoľahlivá premena energie, ale ich hlavnou nevýhodou sú vzájomne prepojené zdroje a záťaže s rôznymi úrovňami napätia.

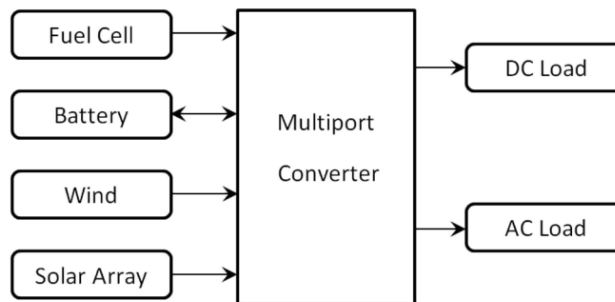
Konvenčnou metódou na prepájanie zdrojov a záťaží je použitie niekoľkých DC-DC alebo DC-AC meničov pripojených na spoločnú jednosmernú zbernicu, čo je znázornené na Obr. 1. Táto topológia vyžaduje viacnásobnú premenu energie, čo spôsobuje nižšiu účinnosť z dôvodu použitia mnohých meničov elektrickej energie. Použitie viacerých meničov tiež zvyšuje náklady na systém a znižuje hustotu výkonu celého systému.



Obr. 1 Konvenčná topológia prepojenia meničov

Ďalším riešením na vzájomné prepojenie zdroja a záťaže je použitie multiport výkonových meničov. Takýto menič je zobrazený na Obr. 2. Tieto meniče môžu transformovať energiu medzi jednotlivými portami v jednom konverznom kroku. Multiport meniče zvyčajne na svoju činnosť používajú jeden transformátor s viacerými vinutiami alebo viac transformátorov s primárnym a sekundárnym vinutím a nemajú spoločnú zbernicu tak, ako ju poznáme z prepojenia viacerých meničov. V aplikáciách, kde je nevyhnutný systém akumulácie energie, musí byť port na

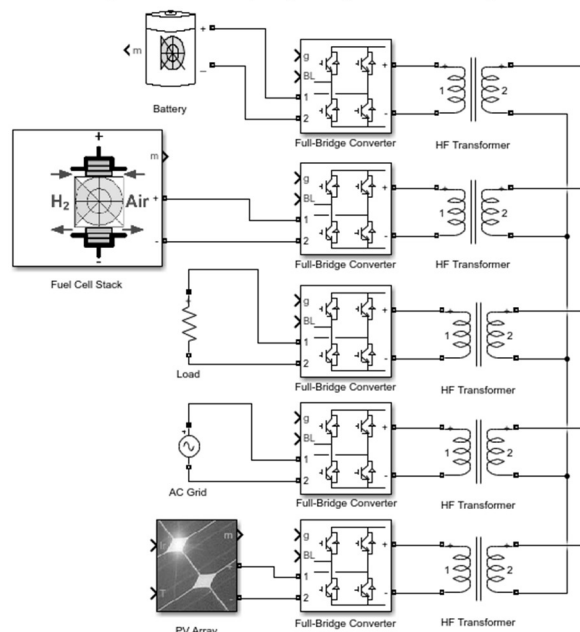
pripojenie tohto akumuláčného systému obojsmerný. Všeobecne sa všetky porty multiport výkonového meniča považujú za obojsmerné. Z tohto dôvodu nie je potrebné rozlišovať medzi vstupným a výstupným portom, preto sa tento menič nazýva multiport menič namiesto viacvstupový alebo viacvýstupový menič. Hlavnými výhodami sú menšie rozmery celého systému, lepšia hustota výkonu, menší počet krokov konverzie a pre multiport menič s viacerými transformátormi s primárnym a sekundárnym vinutím je to tiež vysoká modularita meniča.



Obr. 2 Topológia multiport meniča

## II. MULTIPORT VÝKONOVÝ MENIČ

Navrhovaný multiport menič je určený pre aplikácie, kde sa vyžaduje vysoká modularita, napr. elektrické vozidlá a obnoviteľné zdroje energie. To znamená, že hlavné parametre meniča, ako je počet portov, výkon a typ zdroja a záťaž pripojené k meniču, je možné jednoducho upraviť. Transformátory s viacerými vinutiami sa zvyčajne navrhujú a používajú na konkrétnu aplikáciu s definovaným počtom vinutí. Pridanie alebo odstránenie vinutí z takéhoto transformátora je preto komplikované. Jadro transformátora musí byť navrhnuté na maximálny výkon všetkých zdrojov, pretože viac vinutí na pôvodnom jadre môže naraziť buď na geometrické, alebo magnetické limity jadra. Po pridaní vinutia sa môže jadro presycovať. Aby sa zabránilo problémom s limitmi spoločného jadra transformátora v meniči, navrhovaný menič používa niekoľko transformátorov s paralelne spojenými sekundárnymi vinutiami. Rozšírenie meniča je potom jednoduché pridaním ďalšieho transformátora s potrebnými parametrami. Takýto prístup ponúka vyššiu modularitu meniča. Prenos energie medzi jednotlivými portami je potom riadený pomocou plného mostíka pripojeného na primárnu stranu každého transformátora tak, ako je zobrazené na Obr. 3. V porovnaní s konvenčnou topológiou je počet konverzií energie znížený a je zachovaná vysoká modularita. Modularita, nízka hmotnosť a malé rozmery meniča je možné zachovať aj vďaka použitiu vysokofrekvenčného planárneho transformátora. Tieto transformátory s jadrom EI majú nízke straty, vynikajúci odvod tepla a malý objem [1], [2].

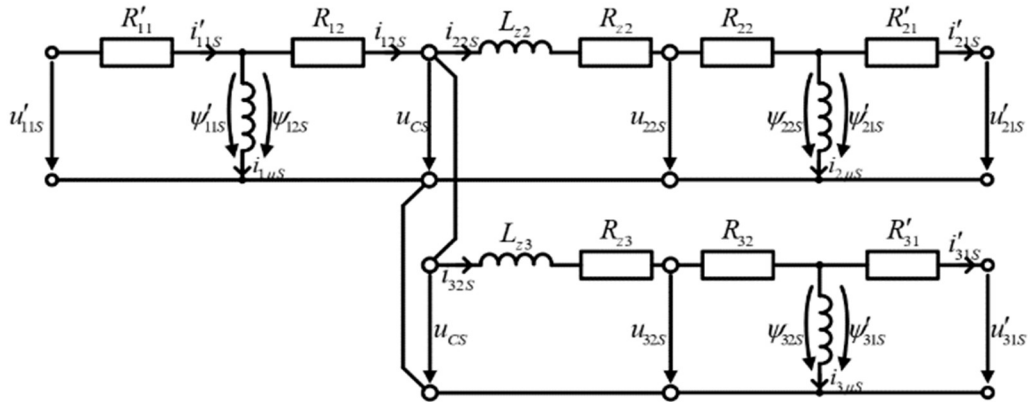


Obr. 3 Multiport výkonový menič s mostíkmi na primárnej strane

## III. MATEMATICKÝ MODEL VIACPORTOVÉHO VÝKONOVÉ MENIČA

Popis multiport meniča je založený na teórii elektrických strojov. Náhradná schéma multiport meniča je na Obr. 4. Všetky parametre v náhradnej schéme sú vzťahnuté na sekundárnu stranu

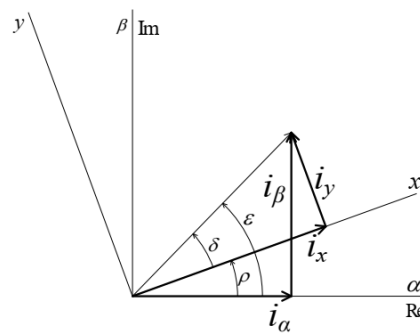
transformátora. Parametre primárnych vinutí transformátorov sú teda prepočítané na sekundárnu stranu a označené apostroфом.



Obr. 4 Náhradná schéma multiport meniča

Na popis transformátora použijeme súradnicový systém zobrazený na Obr. 5. Veličiny na obrázku sú striedavé, časovo premenné veličiny. Ak ich umiestnime do komplexnej roviny a označíme reálnu os ako  $\alpha$  a imaginárnu os ako  $\beta$ , bude rovnica vstupného napätia:

$$\mathbf{u}_{11S}(t) = u_{11\alpha}(t) + ju_{11\beta}(t) = u_{11}(t)e^{j\varepsilon}. \quad (1)$$



Obr. 5 Súradnicový systém použitý pri popise transformátorov

Súradnicový systém s osami  $\alpha$  a  $\beta$  budeme nazývať stojací systém. Veličiny v tomto súradnicovom systéme budú označené indexom „S“. Súradnicový systém s osami  $x$ ,  $y$  označíme ako točivý systém. Tento systém rotuje určitou uhlovou rýchlosťou. Premenné v tomto systéme budú označené indexom „K“. Rovnica, ktorá popisuje vstupné napätie v točivom systéme:

$$\mathbf{u}_{11K}(t) = u_{11x}(t) + ju_{11y}(t) = u_{11}(t)e^{j\delta}. \quad (2)$$

Označenie závislosti veličín na čase ďalej nebudeme kvôli prehľadnosti uvádzať. Rovnica popisujúca transformáciu medzi stojacim a točivým systémom:

$$\frac{\mathbf{u}_{11S}}{\mathbf{u}_{11K}} = \frac{u_{11}e^{j\varepsilon}}{u_{11}e^{j\delta}} = e^{j(\varepsilon-\delta)} = e^{j\rho}. \quad (3)$$

Rovnica pre transformáciu zo stojaceho do točivého systému je:

$$\mathbf{u}_{11S} = \mathbf{u}_{11K}e^{j\rho}, \quad (4)$$

a naopak

$$\mathbf{u}_{11K} = \mathbf{u}_{11S}e^{-j\rho}. \quad (5)$$

Matematický model bol odvodený za nasledujúcich predpokladov:

- Veličiny a parametre na primárnej strane transformátorov sú prepočítané na sekundárnu stranu.

- Sekundárne vinutia transformátorov sú prepojené paralelne. Toto spojenie budeme nazývať ako medziobvod.
- Systém je napájaný harmonickým napätím.
- Rozptylové indukčnosti transformátorov predstavujú približne 1% z celkovej indukčnosti transformátorov. V náhradnej schéme sa preto neobjavujú a sú zahrnuté do indukčností v medziobvode.
- Budeme predpokladať prepojenie troch transformátorov.

Pre primárny obvod transformátora  $T_1$  platia rovnice:

$$\mathbf{u}'_{11S} = R'_{11}\mathbf{i}'_{11S} + \frac{d\psi'_{11S}}{dt}, \quad (6)$$

$$\mathbf{i}'_{11S} = \mathbf{i}_{12S} + \mathbf{i}_{1\mu S}, \quad (7)$$

$$\psi'_{11S} = L'_{11h}(\mathbf{i}'_{11S} - \mathbf{i}_{12S}) = L'_{11h}\mathbf{i}_{1\mu S} = L_{12h}\mathbf{i}_{1\mu S}, \quad (8)$$

$$\frac{d\psi'_{11S}}{dt} = \frac{d\psi_{12S}}{dt} = \mathbf{u}'_{i11S} = \mathbf{u}_{i12S}. \quad (9)$$

Prevod transformátora  $T_1$  je definovaný ako:

$$\frac{u_{i11S}}{u_{i12S}} = r_1, \quad \text{potom} \quad \mathbf{u}'_{i11S} = \frac{u_{i11S}}{r_1}. \quad (10)$$

Podobne platí pre prúdy:

$$\frac{i_{11S}}{i_{12S}} = \frac{1}{r_1}, \quad \text{potom} \quad i'_{i11S} = r_1 i_{11S}. \quad (11)$$

Podobné rovnice platia aj pre transformátory  $T_2$  a  $T_3$ . Pre medziobvod platia nasledujúce rovnice:

$$\mathbf{i}_{12S} = \mathbf{i}_{22S} + \mathbf{i}_{32S}. \quad (12)$$

Pre vetvu medzi  $T_1$  a  $T_2$ :

$$\mathbf{u}_{i12S} = R_{12}\mathbf{i}_{12S} + \mathbf{u}_{CS}. \quad (13)$$

$u_{CS}$  je napätie v bode, kde sú sekundárne vinutia všetkých transformátorov spojené. Pre  $u_{CS}$  preto platí:

$$\mathbf{u}_{CS} = (R_{z2} + R_{22})\mathbf{i}_{22S} + L_{z2} \frac{di_{22S}}{dt} + \mathbf{u}_{i22S}. \quad (14)$$

Ak dosadíme rovnice (12) a (14) do (13)

$$\mathbf{u}_{i12S} = R_{12}\mathbf{i}_{22S} + R_{12}\mathbf{i}_{32S} + R_{z2}\mathbf{i}_{22S} + R_{22}\mathbf{i}_{22S} + L_{z2} \frac{di_{22S}}{dt} + \mathbf{u}_{i22S} \quad (15)$$

a po úprave

$$L_{z2} \frac{di_{22S}}{dt} = \mathbf{u}_{i12S} - \mathbf{u}_{i22S} - R_{12}\mathbf{i}_{32S} - (R_{12} + R_{22} + R_{z2})\mathbf{i}_{22S}. \quad (16)$$

Podobne odvodíme rovnice pre vetvu medzi  $T_1$  a  $T_3$ . Rovnica pre  $u_{CS}$  sa zmení nasledovne:

$$\mathbf{u}_{CS} = (R_{z3} + R_{32})\mathbf{i}_{32S} + L_{z3} \frac{di_{32S}}{dt} + \mathbf{u}_{i32S}. \quad (17)$$

Ak dosadíme (12) a (17) do (13), dostaneme

$$\mathbf{u}_{i12S} = R_{12}\mathbf{i}_{22S} + R_{12}\mathbf{i}_{32S} + R_{23}\mathbf{i}_{32S} + R_{33}\mathbf{i}_{32S} + L_{23}\frac{di_{32S}}{dt} + \mathbf{u}_{i32S} \quad (18)$$

a po úprave

$$L_{23}\frac{di_{32S}}{dt} = \mathbf{u}_{i12S} - \mathbf{u}_{i32S} - R_{12}\mathbf{i}_{22S} - (R_{12} + R_{32} + R_{23})\mathbf{i}_{32S}. \quad (19)$$

#### A. Transformácia do $\alpha$ - $\beta$ Systému

Výsledná podoba rovníc je založená na veličinách, ktoré chceme riadiť. Princíp riadenia systému bude nasledovný:

- Systém umožňuje nezávislé riadenie činnnej a jalovej zložky prúdu.
- Jalové zložky sú primárne riadené portom 1. Výstupné napätie  $T_1$  je riadené magnetizačným prúdom  $i_{1\mu}$ . Pretože odpor a rozptylová indukčnosť transformátorov sú malé, je rozdiel medzi indukovaným napätím a koncovým napätím na primárnej a sekundárnej strane malý. Indukované napätie transformátora  $T_1$  definuje napätie v medziobvode  $u_C$ . Prúd  $i_{1\mu}$  bude riadený napätím  $u'_{11}$ . Magnetizačné prúdy pre port 2 a 3 vytvárajú regulátory prúdov  $i_{22}$  a  $i_{32}$  podľa napätia  $u_C$ .
- Činné a jalové prúdy portu 2 a 3 sú riadené regulátormi prúdov  $i_{22}$ , respektíve  $i_{32}$ . Tieto prúdy sú riadené napätiami  $u'_{21}$  respektíve  $u'_{31}$ . Regulátory prúdov  $i_{22}$  a  $i_{32}$  nastavujú činný prúd, ktorý porty dodávajú alebo odoberajú z medziobvodu. Rozdiel medzi oboma aktívnymi prúdmi dodáva port 1.

Ďalej bude opísaný postup transformácie primárneho obvodu  $T_1$  do točivého systému. Pri transformácii vychádzame z rovníc (6), (7) a (8). Rovnice upravíme tak, aby sme vyjadrili prúd  $i_{1\mu}$

$$\mathbf{u}'_{11S} = R'_{11}\mathbf{i}'_{11S} + \frac{d\psi'_{11S}}{dt} = R'_{11}\mathbf{i}_{1\mu S} + R'_{11}\mathbf{i}_{12S} + L_{12h}\frac{di_{1\mu S}}{dt}, \quad (20)$$

a na rovnicu aplikujeme (4)

$$\begin{aligned} \mathbf{u}'_{11S} = \mathbf{u}'_{11K}e^{j\rho} = R'_{11}\mathbf{i}'_{11K}e^{j\rho} + \frac{d}{dt}(\psi'_{11K}e^{j\rho}) = R'_{11}\mathbf{i}_{1\mu K}e^{j\rho} + R'_{11}\mathbf{i}_{12K}e^{j\rho} + \\ + L_{12h}\frac{d}{dt}(\mathbf{i}_{1\mu K}e^{j\rho}) \end{aligned} \quad (21)$$

Transformačný uhol  $\rho$  je integrálom otáčavej rýchlosti  $\omega_K$ . Táto rýchlosť je daná frekvenciou striedavého napätia, ktorým sú transformátory napájané a je rovnaká pre všetky transformátory. Deriváciu  $i_{1\mu}$  potom odvodíme ako:

$$\frac{d}{dt}(\mathbf{i}_{1\mu K}e^{j\rho}) = \frac{di_{1\mu K}}{dt}e^{j\rho} + j\frac{d\rho}{dt}\mathbf{i}_{1\mu K}e^{j\rho} = \frac{di_{1\mu K}}{dt}e^{j\rho} + j\omega_K\mathbf{i}_{1\mu K}e^{j\rho}. \quad (22)$$

Ak dosadíme (22) do (21) a upravíme, dostaneme:

$$\mathbf{u}'_{11S} = R'_{11}\mathbf{i}_{1\mu K}e^{j\rho} + R'_{11}\mathbf{i}_{12K}e^{j\rho} + L_{12h}\frac{di_{1\mu K}}{dt}e^{j\rho} + j\omega_K L_{12h}\mathbf{i}_{1\mu K}e^{j\rho}. \quad (23)$$

Ak rovnicu (23) vynásobíme  $e^{-j\rho}$  a uvažime

$$\mathbf{u}'_{11K} = \mathbf{u}'_{11S}e^{-j\rho}, \quad (24)$$

môžeme rovnicu upraviť do tvaru

$$L_{12h}\frac{di_{1\mu K}}{dt} = \mathbf{u}'_{11K} - R'_{11}\mathbf{i}_{1\mu K} - R'_{11}\mathbf{i}_{12K} - j\omega_K L_{12h}\mathbf{i}_{1\mu K}, \quad (25)$$

respektíve

$$\frac{di_{1\mu K}}{dt} = \frac{1}{L_{12h}}\mathbf{u}'_{11K} - \frac{R'_{11}}{L_{12h}}\mathbf{i}_{1\mu K} - \frac{R'_{11}}{L_{12h}}\mathbf{i}_{12K} - j\omega_K\mathbf{i}_{1\mu K}. \quad (26)$$

Rovnicu uvedenú vyššie môžeme rozpísať ako

$$\frac{d}{dt}(i_{1\mu x} + ji_{1\mu y}) = \frac{1}{L_{12h}}(u'_{11x} + ju'_{11y}) - \frac{R'_{11}}{L_{12h}}(i_{1\mu x} + ji_{1\mu y}) - \frac{R'_{11}}{L_{12h}}(i_{12x} + ji_{12y}) - j\omega_K(i_{1\mu x} + ji_{1\mu y}), \quad (27)$$

a ďalej rozdeliť (27) na reálnu a imaginárnu časť:

$$\frac{di_{1\mu x}}{dt} = \frac{1}{L_{12h}}u'_{11x} - \frac{R'_{11}}{L_{12h}}i_{1\mu x} - \frac{R'_{11}}{L_{12h}}i_{12x} + \omega_K i_{1\mu y}, \quad (28)$$

a

$$\frac{di_{1\mu y}}{dt} = \frac{1}{L_{12h}}u'_{11y} - \frac{R'_{11}}{L_{12h}}i_{1\mu y} - \frac{R'_{11}}{L_{12h}}i_{12y} - \omega_K i_{1\mu x}. \quad (29)$$

Pre komponenty indukovaného napätia  $u_{i12K}$  potom platia rovnice:

$$u_{i12x} = L_{12h} \frac{di_{1\mu x}}{dt} - \omega_K L_{12h} i_{1\mu y}, \quad (30)$$

a

$$u_{i12y} = L_{12h} \frac{di_{1\mu y}}{dt} + \omega_K L_{12h} i_{1\mu x}. \quad (31)$$

Podobne ako pri transformácii pre primárny obvod transformátora  $T_1$  zavedieme tiež rovnice pre primárny obvod transformátora  $T_2$ , transformátora  $T_3$  a medziobvod. Pre rozsiahlosť matematického aparátu nebudeme ďalej uvádzať postup pre transformácie  $T_2$ ,  $T_3$  a medziobvod. Uvedieme iba konečné rovnice pre prúd a napätie rozdelené na reálnu a imaginárnu časť.

Pre transformátor  $T_2$  platí:

$$\frac{di_{2\mu x}}{dt} = \frac{1}{L_{22h}}u'_{21x} - \frac{R'_{21}}{L_{22h}}i_{2\mu x} + \frac{R'_{21}}{L_{22h}}i_{22x} + \omega_K i_{2\mu y}, \quad (32)$$

$$\frac{di_{2\mu y}}{dt} = \frac{1}{L_{22h}}u'_{21y} - \frac{R'_{21}}{L_{22h}}i_{2\mu y} + \frac{R'_{21}}{L_{22h}}i_{22y} - \omega_K i_{2\mu x}. \quad (33)$$

Indukované napätie pre  $T_2$ :

$$u_{i2x} = L_{22h} \frac{di_{2\mu x}}{dt} - \omega_K L_{22h} i_{2\mu y}, \quad (34)$$

$$u_{i2y} = L_{22h} \frac{di_{2\mu y}}{dt} + \omega_K L_{22h} i_{2\mu x}. \quad (35)$$

Pre transformátor  $T_3$ :

$$\frac{di_{3\mu x}}{dt} = \frac{1}{L_{32h}}u'_{31x} - \frac{R'_{31}}{L_{32h}}i_{3\mu x} + \frac{R'_{31}}{L_{32h}}i_{32x} + \omega_K i_{3\mu y}, \quad (36)$$

$$\frac{di_{3\mu y}}{dt} = \frac{1}{L_{32h}}u'_{31y} - \frac{R'_{31}}{L_{32h}}i_{3\mu y} + \frac{R'_{31}}{L_{32h}}i_{32y} - \omega_K i_{3\mu x}. \quad (37)$$

Indukované napätie pre  $T_3$ :

$$u_{i3x} = L_{32h} \frac{di_{3\mu x}}{dt} - \omega_K L_{32h} i_{3\mu y}, \quad (38)$$

$$u_{i3y} = L_{32h} \frac{di_{3\mu y}}{dt} + \omega_K L_{32h} i_{3\mu x}. \quad (39)$$

Pre medziobvod platia rovnice:

- pre vetvu medzi  $T_1$  a  $T_2$ :



$$L_{z2} \frac{di_{22x}}{dt} = u_{i12x} - u'_{21x} - R'_{21} i'_{21x} - R_{12} i_{32x} - R_{c2} i_{22x} + \omega_K L_{z2} i_{22y}, \quad (40)$$

$$L_{z2} \frac{di_{22y}}{dt} = u_{i12y} - u'_{21y} - R'_{21} i'_{21y} - R_{12} i_{32y} - R_{c2} i_{22y} - \omega_K L_{z2} i_{22x}. \quad (41)$$

- pre vetvu medzi  $T_1$  a  $T_3$ :

$$L_{z3} \frac{di_{32x}}{dt} = u_{i12x} - u'_{31x} - R'_{31} i'_{31x} - R_{12} i_{22x} - R_{c3} i_{32x} + \omega_K L_{z3} i_{32y}, \quad (42)$$

$$L_{z3} \frac{di_{32y}}{dt} = u_{i12y} - u'_{31y} - R'_{31} i'_{31y} - R_{12} i_{22y} - R_{c3} i_{32y} - \omega_K L_{z3} i_{32x}. \quad (43)$$

Kompletný matematický model multiport meniča s tromi portami je teda popísaný rovnicami (28) a (29) pre  $i_{1\mu K}$ , (32) a (33) pre  $i_{2\mu K}$ , (36) a (37) pre  $i_{3\mu K}$ , (40) a (41) pre  $i_{22K}$  a (42) a (43) pre  $i_{32K}$ .

Pre takto popísaný matematický model je ďalej možné navrhnuť riadenie. Ako veličiny, ktoré chceme riadiť, uvažujeme magnetizačný prúd  $i_{1\mu}$  a prúdy  $i_{22}$  a  $i_{32}$ . Pre tieto veličiny budú ďalej navrhnuté PI regulátory, ktorých úloha bude ako nastaviť magnetizačný prúd jednotlivých transformátorov, tak aj nastavovať odoberaný alebo dodávaný prúd cez jednotlivé porty.

#### IV. ZÁVER

Príspevok predstavuje matematický popis topológie výkonového meniča s viacerými portami. Pre popis multiport meniča je využitý matematický aparát známy z teórie elektrických strojov. Multiport výkonový menič danej topológie ponúka vysokú modularitu, nízky počet konverzií a vysokú výkonovú hustotu. Vysokou modularitou sa rozumie najmä jednoduchá zmena počtu portov meniča prostredníctvom prídania ďalšieho transformátora s požadovanými parametrami.

Ďalší výskum bude zameraný na návrh a verifikáciu riadenia meniča, ktorého úlohou bude nastavenie ako magnetizačných prúdov pre jednotlivé transformátory, tak aj nastavenie prúdov jednotlivých portov. Jednotlivé porty multiport meniča budú schopné prúd odoberať aj dodávať.

#### REFERENCES

- [1] TANA, Linlin; YANG, Xu; YUWEN, Dian. Optimization design of planar transformer in DC-DC converter. 2016.
- [2] A. T. Mahamat, B. Danoumbé, M. Youssouf, J. P. Chatelon, S. Capraro and J. J. Rousseau, "Optimization of integrated magnetic planar transformer," 2016 18th European Conference on Power Electronics and Applications (EPE'16 ECCE Europe), Karlsruhe, 2016, pp. 1-8. doi: 10.1109/EPE.2016.7695386.

# Perspectives of OR/TS2JavaConn Utilization in Automated Rail Control Simulation

<sup>1</sup>Tomáš TALIAN, <sup>2</sup>Štefan KOREČKO, <sup>3</sup>Branislav SOBOTA

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

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

**Abstract** – Within the last years, there has been an effort at the home institution of the authors to expand the features of the Open Rails railroad simulator in terms of an automatic operation of trains via the means of external control modules. This work utilizes the said project, the platform OR/TS2JC, to emulate the operation of the European Train Control System using the available methods and resources of the project to their full potential.

**Keywords** – Railroad, Simulation, Automatic Operation, Open Rails, ETCS

## I. INTRODUCTION

Contemporary railroad safety systems are the main contributing factor to making the railway one of the safest modes of transport. These safety systems come in various forms, with different levels of complexity, and utilize a range of methods, but their main goal is always the same: to prevent accidents. They underwent a long process of evolution, which still continues today and during which their complexity and reliability gradually increased. Their current state was preceded by many accidents from which valuable data have been collected and used in further development and upgrades to prevent such accidents in the future.

Teams of experts and researchers carry out extensive testing, validation, and simulations before applying changes to existing systems or before a brand new safety system is put into operation. Safety systems can be interconnected with train control systems to achieve varying levels of autonomous operation. With the most advanced and secure systems, even a fully automatic operation is possible. These solutions rely heavily on maximum reliability, error prevention, and safety.

One of the ways of simulating rail traffic is using already existing simulator computer games. The platform OR/TS2JC [1], [2], [3] has been developed at the home institution of the authors by modifying the open-source Open Rails (OR) [4] simulator and connecting it with an original application, called Train Simulator to Java Connector (TS2JC). TS2JC allows separately developed control modules to manage a railway scenario, simulated in OR. A control module reacts to events that occur during the simulation.

With the advent of the modern railway control systems, such as the European Train Control System (ETCS), a question has arisen whether the OR/TS2JC platform is capable of emulating their operation via appropriate control programs, called control modules. In this paper we briefly report on an experiment, where a railway scenario modeling a part of a real railway route was selected and a control module emulating the operation of ETCS as closely as possible was selected for it.

## II. RELATED WORK

Using simulation and other types of computer games for research, experiments, training, or education is not a new concept. There are multiple ways of enabling this wider use of games outside of just the pure entertainment. To emphasize this non-standard use, such games are often labeled by the term “games with purpose”, coined by Luis Von Ahn in [5]. One of the ways is to take an open-source game, such as OR, which source code is directly editable. However, very few games offer this and therefore it is necessary to look for other options. Using games which offer an Application Programming Interface (API) is a great way of tapping into their potential and capabilities. Some games even offer tools provided by their developers which allow adding content or modifying parts of the game in some way. One other possibility is to directly

communicate and work with the developer and agree to making a custom version under specific licensing conditions, but this can be overly complicated and unfeasible.

Our work is not the first one to utilize Open Rails. For example, in [6] researchers from the University of Žilina used a control panel extracted from a real locomotive and connected it with the Open Rails simulator. They have created custom software and hardware interfaces, communicating with a console application which in turn communicates with the simulator. The interfaces also allow to connect other types of train hardware.

In [7], a research group from Singapore used OR to tackle the problem of cyber attacks on railroad systems. They added simulation of passenger movements and together with Matlab they were able to simulate different attacks and analyze the disruptions. From the collected data they were able to assess possible effectiveness and feasibility of a proposed cyber attack detection system.

Flight simulators are also a popular research category. In [8], developers from the University of Porto used the API and Software Development Kit (SDK) of Flight Simulator X to create a distributed simulator able to simulate a greater area than previously possible. Using multiple synchronized simulator instances they were able to seamlessly travel through different game areas by passing the planes and data to the respective instances simulating neighboring areas. In a similar way, OR/TS2JC is using multiple OR instances to simulate multiple trains on the same route.

Another example of re-purposing an entertaining game for scientific purposes is Starcraft, which is commonly associated with experiments and research in the area of Artificial Intelligence (AI). As the game is not open-source, the provided API is used. One of the works using Starcraft is [9], where evolution algorithms are utilized to play the game effectively according to various parameters of the Starcraft game mechanic, which are used to create a great variety of different scenarios.

### III. OVERVIEW OF OR/TS2JC

As described in more detail in [2], OR/TS2JC simulation software consists of three main parts:

- customized version of the Open Rails simulator,
- TS2JC application and
- control modules.

Open Rails (OR) has been extended by the ability to connect to an external application. It is able to issue and receive commands, and respond to them according to the current state and events in the simulated scenario. It uses in-game actions and triggers to raise external events, which are communicated to other parts of the platform.

TS2JC is a custom-made graphical application, able to communicate with a running server instance of the extended OR simulator. It serves as an intermediary application and an interface to bridge the simulated scenario and the control modules. It receives all the necessary simulation data and elements such as track sections, trains, signals, switches, and stations, to have an overall overview of the simulation's state. Fig. 1 shows the Graphical User Interface (GUI) of the application, where all elements are divided into their respective sections. The text visible in the logger part contains a portion of the data string received from the simulator which has been parsed into individual elements.

The mentioned data are also used in the control modules. TS2JC contains a built-in generator which is able to create a control module skeleton with all data included and also containing methods to provide some basic functionality. Modules can be generated in the Java language and the languages of formal methods B-Method and Perfect Developer. They can be independently programmed to achieve the desired functionality.

The main operation of OR/TS2JC is based on the events from the simulator using the challenge-response type of communication. All events originate from OR, which notifies the TS2JC application, which in turn notifies the currently active control module. Through a common interface, a corresponding method of the module is executed according to the type of the raised event. Module methods update their own internal simulation states and those changes are then propagated back to the simulator and also visualized in the GUI of TS2JC (Fig. 1).

### IV. EMULATION OF ETCS

To emulate the ETCS, first the system itself had to be analyzed to determine its main operating principles. The current state of OR/TS2JC also had to be analyzed to assess which methods and resources are available to emulate different aspects of the ETCS. The following features of OR/TS2JC were determined to be the main focus of the ETCS emulation effort:

- *Movement Authority* - permission for a train to move to a new location, periodically refreshed depending on the traffic situation.

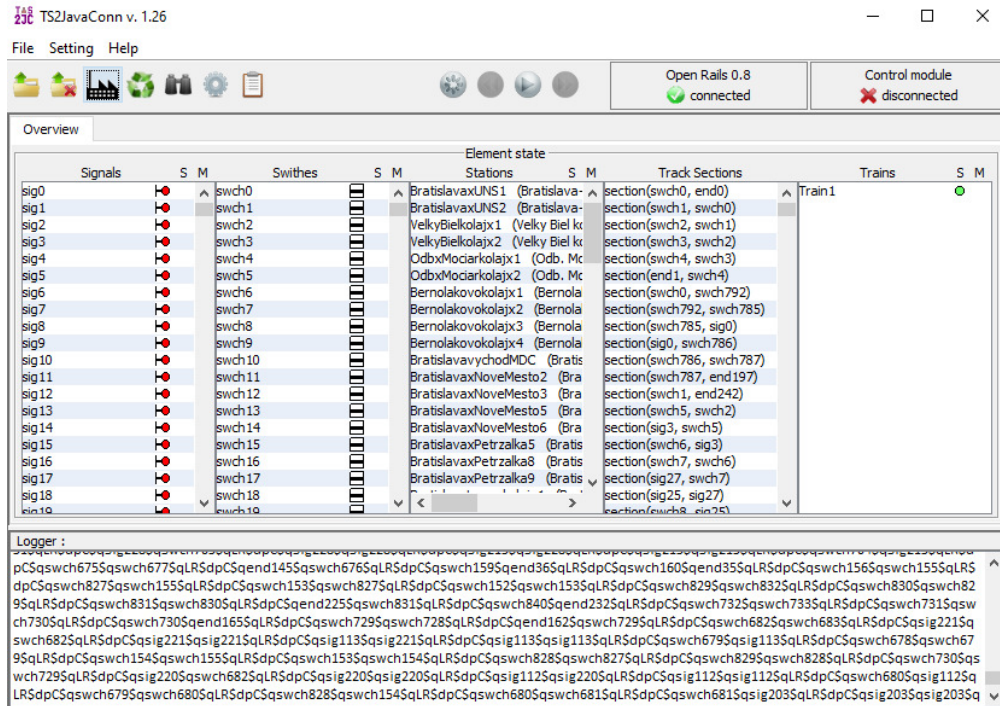


Fig. 1 Graphical user interface of TS2JC.

- *Track sections monitoring* - dividing the tracks into sections and monitoring their occupancy states to determine whether a train should be allowed into the section.
- *Train location monitoring* - knowing the exact location of every train at any moment for use in other parts of the system.
- *Train speed control* - automatic acceleration, braking, cruising, and stopping, according to the speed limits and track conditions.
- *Signaling* - informing the driver about the track situation and what to expect in front of the train.

The next step was a selection of a railway scenario for which an ETCS-emulating control module will be developed. We decided to use an already existing scenario, which models a real route of non-trivial complexity. After testing several scenarios, the scenario [10], which models the route 130 of Slovak railways (ŽSR), has been selected. The OR/TS2JC control modules use enumerated sets to represent track elements, such as signals, switches, sections or stations. To not exceed the maximum number of enumerated elements, allowed by the Java language, it was necessary to reduce the size of the scenario.

The analysis of OR/TS2JC as well as compatibility issues revealed when testing the route 130 scenario resulted in minor implementation changes in OR/TS2J, without any alteration of the simulator behavior. The most important modifications were handling of incorrect track endings and station names in OR scenarios, handling of the *getSpeed* event and a generation of an auxiliary text file to make identification of track elements easier.

The final step was an implementation of the *ETCSlike* control module, which emulates the ETCS operation as close as possible. Each control module contains two types of methods: *getters*, which return values of module variables that correspond to states of the simulated scenario components and *modifiers*, which are called when an event occurs in the simulated scenario in OR. The modifiers may change the values of the module variables. From the several types of events which a control module can handle, the *ETCSlike* module deals with the following ones:

- *sectionEnter* - raised when the front of the train enters a new section.
- *sectionLeave* - raised when the back of the train leaves a section.
- *requestGreen* - raised periodically while the next signal (in front of the train) is red (danger).
- *redSignalApproach* - raised once when a train reaches the predetermined distance from a red signal.
- *signalStateChanged* - raised when the state of the next signal changes (e.g. from red to green).
- *getSpeed* - raised when the simulator requests the current speed of a specific train.

The ETCS uses track devices called Eurobalise and Euroloop, or radio communication, to locate

the trains and deliver crucial information to them. Eurobalises are devices located in the tracks capable of sending information to a passing train. Euroloops are cables, which serve the same purpose as Eurobalises with the difference being that they are able to send information continuously throughout the whole length of the cable. Higher levels of ETCS use radio communication to transmit the information, with Eurobalises and Euroloops being used only for calibration. These devices are not available in OR/TS2J, but we can emulate them using the mentioned events.

The implementation of the *ETCSlike* module started with a skeletal module for the route 130 scenario, created with the built-in module generator of TS2JC. Java has been selected as the programming language. Some of the methods were kept unchanged, other were updated to utilize the ETCS-like event handling.

The OR/TS2JC platform creates track sections, which are the base of basically every railroad safety system, by pairing adjacent track elements, namely switches, signals and track entry points. To make section handling easier, custom grouped sections were created as a part of the *ETCSlike* module. These sections are made of two signals valid for one travel direction and contain a list of all subsections between these signals. Additional information, such as the occupancy state, the identifier of the train occupying the section and the section speed, are being monitored to satisfy some of the aspects of the ETCS mentioned above.

A representation of trains and their states has been implemented into the *ETCSlike* module. These train objects keep information about their route and about the proper switch configurations to prevent them entering a section outside of their defined routes. They keep their current location and their current maximum speed, which updates regularly in case it is requested by the simulator.

The task of issuing the permissions for train movement has been delegated to a method emulating the Movement Authority. Every time a train enters or leaves a section, or requests a permission to move, the Movement Authority is actualized and, according to the traffic situation, is either extended or left in its current state. When the Movement Authority is issued for a certain distance, represented by a maximum number of sections, all relevant track sections are automatically marked as occupied by the train to which the authority was issued. All switches contained in these sections are set to their desired positions, according to the train route and switch configurations.

Fig. 2 shows a model situation of one train following another under the supervision of the *ETCSlike* module. The pink and blue colored lines depict the Movement Authority for the trains A and B respectively. Black lines represent unoccupied sections. Multiple track sections can be marked as occupied by a train even before entering them. Three signal states can be seen, proceed (green), caution (yellow), and danger (red), which are updated every time a train leaves or enters a section, or a new Movement Authority is issued. Signal states are always stored in the module as well as visualized in TS2JC.

Our module has been tested in various situations that can occur in everyday railroad traffic with a single or multiple trains at once. These situations include but are not limited to:

- trying to enter an occupied section,
- waiting for a section to free,
- crossing of the train paths,
- changing the speed limits,
- following another train and
- setting a correct train route.

For a correct operation, the module has to be manually configured for a specific route. That means creating the grouped sections which are going to be used during the simulation as well as configuring the train routes and the correct switch positions.

## V. CONCLUSION

The goal of this work was to determine whether the OR/TS2JC simulation software, developed at the Technical University of Košice, could be used to emulate the ETCS on a randomly selected route using the available methods and resources. Both the platform and the safety system have been analyzed to choose an appropriate approach.

After minor changes to OR/TS2JC, which ensure better compatibility with community-created routes and a better use of available methods, it was possible to emulate the ETCS functionality to a certain degree. The functionality was implemented in a control module, called *ETCSlike*, for a part of an already existing railway scenario, modeling the route 130 of Slovak railways. The module was programmed using the Java language.

By the “certain degree” we mean that, thanks to the freedom given to a programmer when defining the internal functionality of control modules, it was possible to implement the decision making as specified in ETCS. This includes the Movement Authority, speed control, train location

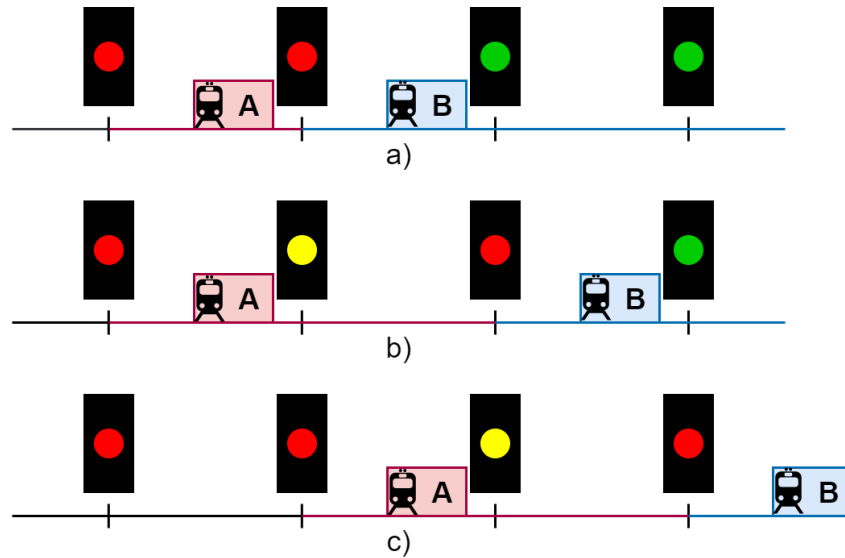


Fig. 2 Model situation under the supervision of a safety system as implemented in the *ETCSlike* control module.

monitoring, track section creation and monitoring, signaling, and train route management. On the other hand, the OR/TS2JC's functionality would have to be altered to provide additional information to the modules, mainly regarding constant and precise updates of train distances to various track elements and other trains. This information could be used to more accurately react to situations on the tracks and help with more precise control of the trains. OR should also be further modified to allow sections defined outside of the traditional track elements (signals and switches). In addition, a more detailed and specialized GUI could be created for the TS2JC to show the various aspects of operation of the emulated system.

#### ACKNOWLEDGMENT

This paper is the impact within the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology II. phase, ITMS: 313011D232, supported by the Research & Development Operational Programme funded by the ERDF.

#### REFERENCES

- [1] Š. Korečko and M. Janotka, "Train simulator as an environment for development of control programs," in *Electrical Engineering and Informatics VI*. Faculty of Electrical Engineering and Informatics, Technical University of Košice, 2015, pp. 849–852.
- [2] Š. Korečko and B. Sobota, "Computer games as virtual environments for safety-critical software validation," *Journal of Information and Organizational Sciences*, vol. 41, no. 2, pp. 197–212, 2017.
- [3] Š. Korečko, T. Hrušč, and S. Janotka, "Or/ts2jc platform as a virtual environment for validation of safety-critical software," in *Electrical Engineering and Informatics X*. Faculty of Electrical Engineering and Informatics, Technical University of Košice, 2019, pp. 280–284, (in Slovak).
- [4] Open Rails Team. Open rails - free train simulator project. [Online]. Available: <http://openrails.org/>
- [5] L. Von Ahn, "Games with a purpose," *Computer*, vol. 39, no. 6, pp. 92–94, 2006.
- [6] P. Sovička, M. Pácha, and P. Rafajdus, "Development of an advanced locomotive simulator," *2017 6th International Youth Conference on Energy (IYCE)*, pp. 1–5, 2017.
- [7] Z.-T. Teo, B. A. N. Tran, S. Lakshminarayana, W. G. Temple, B. Chen, R. P. A. Tan, and D. Yau, "Securerails: Towards an open simulation platform for analyzing cyber-physical attacks in railways," *2016 IEEE Region 10 Conference (TENCON)*, pp. 95–98, 2016.
- [8] C. Rodrigues, D. C. Silva, R. Rossetti, and E. Oliveira, "Distributed flight simulation environment using flight simulator x," *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1–5, 2015.
- [9] N. Othman, J. Decraene, W. Cai, N. Hu, M. Low, and A. Gouaillard, "Simulation-based optimization of starcraft tactical ai through evolutionary computation," *2012 IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 394–401, 2012.
- [10] XTrain. Route 130 of žsr. [Online]. Available: [http://xtrain.alsat.sk/msts/routes/zsr\\_130.htm](http://xtrain.alsat.sk/msts/routes/zsr_130.htm)

# Author's index

- A**  
Andraš Miroslav 171  
Andrejčík Samuel 124, 133
- B**  
Babič František 57, 61, 239, 307  
Baláž Anton 24, 45, 160, 165  
Baumgartner Maroš 285, 422  
Berka Marek 291  
Biňas Miroslav 250  
Bodnár Dávid 399
- C**  
Cimbala Roman 7, 65, 94, 177, 205  
Cymbalák Dávid 269, 274
- Č**  
Čík Ivan 428
- D**  
Dobranská Lenka 307  
Doci Jozef 239  
Drutarovský Miloš 76, 227, 233
- Ď**  
Ďurovský František 382, 399, 458
- E**  
Eliáš Jozef 296
- F**  
Fedor Marek 146  
Fedorišinová Miriama 377  
Florek Daniel 24  
Focko Peter 216  
Fortes Jána 105, 222  
Frivaldský Jozef 210
- G**  
Genči Ján 216  
Girovský Peter 368, 373  
Grešák Marek 200  
Gula Dávid 406
- H**  
Hasin Martin 331  
Havlice Zdeněk 183, 189  
Havran Peter 7, 65, 94, 177, 205, 318  
Havrilla Martin 331  
Herich Dušan 312  
Hrivniak Patrik 70  
Huba Matúš 301  
Hudák Marián 38
- Ch**  
Chlebovec Martin 76
- I**  
Ildža Jákob 244
- J**  
Jakab František 269, 274, 291, 296  
Javorková Sára 38  
Juhár Jozef 285, 422
- K**  
Kačmár Jerguš 195  
Kainz Ondrej 291, 296  
Kanuščák Juraj 352  
Khorshidiyeh Heidar 165  
Kocur Dušan 105, 222  
Kohan Vladimír 94  
Kolárik Michal 417  
Kolcunová Iraida 318  
Kolek Martin 7  
Koniar Lukáš 382  
Korečko Štefan 38, 406, 435, 465  
Korének Martin 274  
Koscelanský Patrik 250  
Križánek Matúš 57  
Kupcová Eva 233  
Kurimský Juraj 318
- L**  
Lapčák Maroš 12, 18, 110, 115, 119, 124, 128, 133, 138, 142
- M**  
Mačák Stanislav 373  
Madoš Branislav 24, 45, 165  
Mach Marián 210, 428  
Majerník Daniel 393  
Maniak Tomáš 348  
Marcinek Adrián 458  
Masrna Matej 435  
Medveď Dušan 70, 82, 88, 171, 200, 244  
Mičko Kristián 61  
Mihalík Ján 100  
Michalénková Lucia 269  
Michalik Peter 442  
Michalko Miroslav 274, 291, 296  
Mondok Filip 160
- N**  
Nezník Richard 154  
Novák Daniel 417  
Novitzká Valerie 51
- O**  
Oliinyk Maksym 65  
Ondáš Stanislav 377  
Ondruško Tomáš 358  
Oravec Jakub 12, 18, 110, 115, 119, 128, 138, 142  
Ovseník Ľuboš 12, 18, 110, 115, 119, 128, 138, 142
- P**  
Papaj Ján 154, 257, 263, 285, 343, 422  
Paralič Ján 301  
Pavlišková Viera 57, 239, 307  
Pavlík Marek 348, 358  
Pella Zuzana 352  
Perduková Daniela 146  
Petrášová Andrea 45  
Petro Viktor 327  
Pietriková Emília 160  
Pleva Matúš 388  
Pomšár Ladislav 442  
Porteleky Tamás 222  
Potočník Bohuslav 183  
Presada Ján 82  
Pynzar Eduard 388
- R**  
Rohač Marek 227
- S**  
Schütz Patrik 279  
Silagyi Pavol 88  
Slinčák Daniel 189  
Sobota Branislav 38, 406, 435, 465  
Stefányi Tara 51  
Steingartner William 195  
Sulír Matúš 393  
Széplaki Richard 189
- Š**  
Šaková Michaela 263  
Šárpataky Ľuboš 205  
Šárpataky Miloš 177, 318  
Švingál Michal 105
- T**  
Talian Tomáš 465  
Turán Ján 124, 133
- V**  
Vaščák Ján 279, 312, 337, 362, 412, 452  
Vašková Gabriela 343  
Vápeník Roman 269  
Veselovská Lucia 183
- Z**  
Záviš Martin 368  
Zbojovský Ján 318  
Zdravecký Norbert 12, 18, 110, 115, 119, 124, 128, 138, 142  
Zolotová Iveta 442  
Zverec Filip 257

**Electrical Engineering and Informatics XII**  
**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  
Edition I, 471 pages  
Number of CD Proceedings: 50 pieces

Editors: prof. Ing. Alena Pietriková, CSc.  
Ing. Emília Pietriková, PhD.

**ISBN 978-80-553-3940-5**